

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2004, 16:59:08 ; Search time 48 Seconds
(without alignments)
1828.662 Million cell updates/sec

Title: US-09-593-793a-113

Perfect score: 2861

Sequence: 1 MVQRLWVSRLLRHRKAQLL.....AIYFATQVDFKSLAKYSA 553

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A Geneseq_19Jun03.*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2861	100.0	553	19	Amino acid encoded
2	2861	100.0	553	19	Prostate tumour sp
3	2861	100.0	553	21	Protein encoded by
4	2861	100.0	553	21	Human immunogenic
5	2861	100.0	553	22	Human prostate CDN
6	2861	100.0	553	22	Human prostate tum
7	2861	100.0	553	22	Human prostate-spe
8	2861	100.0	553	22	Human prostate-spe
9	2861	100.0	553	22	Human P501S invent

10	2861	100.0	553	22	AB74800	Prostate tumour an
11	2861	100.0	553	23	ABG94411	Human prostate tum
12	2861	100.0	553	23	ABG76665	Prostate tumour pr
13	2861	100.0	553	23	AB877575	Human mast cell re
14	2861	100.0	553	23	ABG61900	Prostate cancer-as
15	2861	100.0	553	23	AB895222	Human L1-12 protei
16	2861	100.0	553	23	AAU10324	Human PROST 03. H
17	2861	100.0	553	23	AAU82643	Human breast tumou
18	2861	100.0	553	24	ABU71653	Prostate cancer sp
19	2861	100.0	1079	24	ABU74830	Prostate tumour an
20	2861	100.0	1079	24	ABU71860	Prostate specific
21	2601	90.9	710	23	AAU50661	Thioredoxin-ubiqui
22	2596	90.7	595	22	AAU01318	Alpha prepro-P501S
23	1696	59.3	359	24	ABU71887	Human prostate spe
24	1677.5	58.6	530	23	AAU50662	Thioredoxin-ubiqui
25	1417.5	49.5	371	22	AAU69875	Human prostate CDN
26	1417.5	49.5	371	22	AAU01330	P553S splice varia
27	1417.5	49.5	371	23	AB895335	Human P553S splice
28	1417.5	49.5	371	24	ABU71766	Prostate cancer as
29	1416	49.5	371	22	AAU01362	Human gene 11 enco
30	1416	49.5	371	23	ABG64105	Human albumin fusi
31	1403.5	49.1	400	22	AAU69907	Human prostate pro
32	1403.5	49.1	400	22	AAU01362	Ra12-P501S-E2 cons
33	1403.5	49.1	400	23	AB895367	Ra12-P501S-E2 cons
34	1403.5	49.1	400	24	ABU71798	Prostate cancer as
35	1287	45.0	255	20	AAW85068	Protein encoded by
36	1287	45.0	255	21	AAW85068	Human prostate-rel
37	1287	45.0	255	22	AAU04205	Prostate-specific
38	1287	45.0	255	23	AAU01984	Human prostate-spe
39	1150	40.2	231	21	AAU54369	Amino acid sequenc
40	1120	39.1	252	22	AAU01423	Human secreted pro
41	1014.5	35.5	326	23	AB877571	Human mast cell re
42	519.5	18.2	599	22	ABG60709	Drosophila melanog
43	457.5	16.0	748	22	AAU40327	Human polypeptide
44	426	14.9	123	22	AAU69873	Human prostate CDN
45	426	14.9	123	22	AAU01228	P553S splice varia

ALIGNMENTS

RESULT 1	AAW71869	AAW71869
ID	AAW71869 standard; Protein; 553 AA.	
XX	AAW71869;	
AC		
XX		
DT	06-JAN-1999 (first entry)	
XX		
DE	Amino acid encoded by prostate tumour clone L1-12.	
XX		
KW	Prostate; cancer; tumour; vaccine; immunogen; clone.	
XX		
OS	Homo sapiens.	
XX		
PN	WO9837093-A2.	
XX		
PD	27-AUG-1998.	
XX		
PF	25-FEB-1998; 98WO-US03492.	
XX		
PR	09-FEB-1998; 98US-0020956.	
PR	25-FEB-1997; 97US-0806099.	
PR	01-AUG-1997; 97US-0904804.	
XX	(CORI-) CORIXA CORP.	
PA	Dillon DC, Xu J;	
PI		
XX		
DR	WFI; 1998-609886/51.	
DR	N-PSDB; AAV61201.	
XX		
PT	Polypeptides comprising immunogenic portions of prostate proteins -	

PT used in a vaccine for the treatment of prostate cancer

XX Example 1; Page 82-84; 130pp; English.

XX The present sequence is an immunogenic portion of a prostate tumour
 CC protein. The immunogen, or the DNA encoding it, can be used as a
 CC vaccine for the treatment of prostate cancer. The immunogen was
 CC isolated from a prostate tumour cDNA library obtained by subtracting
 CC a prostate tumour cDNA expression library with a normal tissue cDNA
 CC library.

XX Sequence 553 AA;

Query Match 100.0%; Score 2861; DB 19; Length 553;
 Best Local Similarity 100.0%; Pred. No. 2.2e-269;
 Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVQRLWVSRLLRHRKAQLLLVNLTFGLVCLAAAGITVVPDLLLEVGVEEKFMTWLVIG 60
 Db 1 MVQRLWVSRLLRHRKAQLLLVNLTFGLVCLAAAGITVVPDLLLEVGVEEKFMTWLVIG 60
 Qy 61 PVGLVVCVPLLGASDHWGRGRRRPFITWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
 Db 61 PVGLVVCVPLLGASDHWGRGRRRPFITWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
 Qy 121 ELALLILGVGLLDFCGQVCFPTLEALLSDLPDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
 Db 121 ELALLILGVGLLDFCGQVCFPTLEALLSDLPDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
 Qy 181 IDWDTSAALPYLGTQECFLGILLTLFTLTCVAATLLVAEEAALGPTPEAGLSAPSLPH 240
 Db 181 IDWDTSAALPYLGTQECFLGILLTLFTLTCVAATLLVAEEAALGPTPEAGLSAPSLPH 240
 Qy 241 CCPCRLAFRNLAGLIPRLHQLCCMRPTRLRLFLVAELCSWMLMTFTLYTDFVGEGL 300
 Db 241 CCPCRLAFRNLAGLIPRLHQLCCMRPTRLRLFLVAELCSWMLMTFTLYTDFVGEGL 300
 Qy 301 YQGVPRAPGTEARRHYDEGVNMGSLGLFLOCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
 Db 301 YQGVPRAPGTEARRHYDEGVNMGSLGLFLOCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
 Qy 361 AFPVAAGATCLSHSVAVVTASAALTGTFTSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 Db 361 AFPVAAGATCLSHSVAVVTASAALTGTFTSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 Qy 421 ASSEDSLMTSFLPGPKFGAPPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGTEA 480
 Db 421 ASSEDSLMTSFLPGPKFGAPPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGTEA 480
 Qy 481 RVVPRGICLDLALDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAIYPATQ 540
 Db 481 RVVPRGICLDLALDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAIYPATQ 540
 Qy 541 VFVSKSLAKYSA 553
 Db 541 VFVSKSLAKYSA 553

RESULT 2

AAW69385

ID AAW69385 standard; Protein; 553 AA.

XX AC AAW69385;

DT 25-MAR-2003 (updated)

DT 08-DEC-1998 (first entry)

XX Prostate tumour specific gene clone L1-12 protein.

XX Prostate tumour specific gene; human; prostate cancer; detection;

KW therapy.

XX Homo sapiens.

OS

XX WO9837418-A2.
 XX 27-AUG-1998.
 XX 25-FEB-1998; 98WG-US03690.
 XX 25-FEB-1997; 97US-0806596.
 PR 01-AUG-1997; 97US-0904809.
 PR 09-FEB-1998; 98US-0020747.
 XX (CORI-) CORIXA CORP.
 XX Dillon DC, Xu J;
 XX WPI; 1998-480805/41.
 DR N-PSDB; AAV58586.

XX Novel human prostate specific tumour protein and fragments - useful
 PT for detecting and treating prostate cancers
 XX Example 1; Page 87-89; 141pp; English.

XX This sequence is encoded by a human prostate tumour specific gene, and
 CC can be used in the method of the invention. The method is for detecting
 CC prostate cancer comprises contacting a biological sample with an agent
 CC able to bind an immunogenic portion of a prostate protein (such as
 CC this protein sequence). An antibody which binds to an immunogenic
 CC portion of the prostate protein, and the method can be used to detect,
 CC monitor progression of, or treat prostate cancers. The antibody may
 CC also be conjugated to a therapeutic agent for use in therapy of prostate
 CC cancers.

CC (Updated on 25-MAR-2003 to correct PR field.)

XX Sequence 553 AA;

Query Match 100.0%; Score 2861; DB 19; Length 553;
 Best Local Similarity 100.0%; Pred. No. 2.2e-269;
 Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVQRLWVSRLLRHRKAQLLLVNLTFGLVCLAAAGITVVPDLLLEVGVEEKFMTWLVIG 60
 Db 1 MVQRLWVSRLLRHRKAQLLLVNLTFGLVCLAAAGITVVPDLLLEVGVEEKFMTWLVIG 60
 Qy 61 PVGLVVCVPLLGASDHWGRGRRRPFITWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
 Db 61 PVGLVVCVPLLGASDHWGRGRRRPFITWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
 Qy 121 ELALLILGVGLLDFCGQVCFPTLEALLSDLPDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
 Db 121 ELALLILGVGLLDFCGQVCFPTLEALLSDLPDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
 Qy 181 IDWDTSAALPYLGTQECFLGILLTLFTLTCVAATLLVAEEAALGPTPEAGLSAPSLPH 240
 Db 181 IDWDTSAALPYLGTQECFLGILLTLFTLTCVAATLLVAEEAALGPTPEAGLSAPSLPH 240
 Qy 241 CCPCRLAFRNLAGLIPRLHQLCCMRPTRLRLFLVAELCSWMLMTFTLYTDFVGEGL 300
 Db 241 CCPCRLAFRNLAGLIPRLHQLCCMRPTRLRLFLVAELCSWMLMTFTLYTDFVGEGL 300
 Qy 301 YQGVPRAPGTEARRHYDEGVNMGSLGLFLOCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
 Db 301 YQGVPRAPGTEARRHYDEGVNMGSLGLFLOCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
 Qy 361 AFPVAAGATCLSHSVAVVTASAALTGTFTSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 Db 361 AFPVAAGATCLSHSVAVVTASAALTGTFTSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 Qy 421 ASSEDSLMTSFLPGPKFGAPPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGTEA 480
 Db 421 ASSEDSLMTSFLPGPKFGAPPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGTEA 480
 Qy 481 RVVPRGICLDLALDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAIYPATQ 540

```

Db 481 RVVPGRGICLDLALDLSAPLISQVAPSLFMGSIVQLSQSVTAYMVSAAGLGLVAIFYATQ 540
Qy 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

RESULT 3
ID AAB28527
XX AAB28527 standard; Protein; 553 AA.
AC AAB28527;
DT 07-FEB-2001 (first entry)
DE Protein encoded by human breast tumour cDNA clone P501S.
XX Human; breast tumour antigen; cytostatic; immunotherapy;
KW breast cancer; vaccine.
XX Homo sapiens.
XX WO2000061756-A2.
XX 19-OCT-2000.
XX 10-APR-2000; 2000WO-US09688.
XX 09-APR-1999; 99US-0288950.
XX 02-JUL-1999; 99US-0346327.
XX (CORI-) CORIXA CORP.
XX Reed SG, Xu J, Dillon DC;
XX WPI; 2000-638568/61.
XX N-PSDB; AAC79473.
XX A novel isolated polypeptide comprising an immunogenic portion of a
XX breast cancer protein useful in the detection and treatment of breast
XX cancer -
XX Claim 2; Page 92-93; 95pp; English.
XX The present sequence is encoded by a cDNA sequence which was isolated
XX from a breast tumour cDNA library. It is provided in a specification
XX relating to compounds for immunotherapy and diagnosis of breast cancer.
XX Breast tumour antigens and the polynucleotides that encode them may be
XX used in the production of a pharmaceutical composition to be used in the
XX treatment of breast cancer. Proliferated T cells and incubated antigen
XX presenting cells are also required. The polypeptides and polynucleotides
XX may also be used to produce a vaccine.
XX Sequence 553 AA;
XX Query Match 100.0%; Score 2861; DB 21; Length 553;
XX Best Local Similarity 100.0%; Pred. No. 2.2e-269; Gaps 0;
XX Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MVQRLWVSRLLRHKAQILLVNLTLFGLEVCLAAGITVVPPLLEVGVEKFTMTVLGIG 60
Db 1 MVQRLWVSRLLRHKAQILLVNLTLFGLEVCLAAGITVVPPLLEVGVEKFTMTVLGIG 60
Qy 61 PVLGLVCVPLIGSADHWGRGYRRRPFIFWALSIGILLSLFLIPRAGWLAGLLCPDRPL 120
Db 61 PVLGLVCVPLIGSADHWGRGYRRRPFIFWALSIGILLSLFLIPRAGWLAGLLCPDRPL 120
Qy 121 ELALLILGVGLDPCGVCFPTPEALLSDLPDRDPHCRQAYSVYAFMISLGGCLGYLLPA 180
Db 121 ELALLILGVGLDPCGVCFPTPEALLSDLPDRDPHCRQAYSVYAFMISLGGCLGYLLPA 180
Qy 181 IDWDTLSALAPYLGTEECLEFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPSLSPH 240

```

```

Db 181 IDWDTLSALAPYLGTEECLEFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPSLSPH 240
Qy 241 CCPCARLAFRNLCALLPRHLQCCMRPRTLRLRFVAELCSWMLMTFTFLFYTDFVGEGL 300
Db 241 CCPCARLAFRNLCALLPRHLQCCMRPRTLRLRFVAELCSWMLMTFTFLFYTDFVGEGL 300
Qy 301 YQGVPRAEPTGEARRHYDEGRMGSLGLFLQCAISLVFSLVMDRLVQRFGTFRVYLASVA 360
Db 301 YQGVPRAEPTGEARRHYDEGRMGSLGLFLQCAISLVFSLVMDRLVQRFGTFRVYLASVA 360
Qy 361 APFVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLYHREKQVLPKRYGTGG 420
Db 361 APFVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLYHREKQVLPKRYGTGG 420
Qy 421 ASSEDSLMTSFLPGPKPGAPFPNGHVGGAGSGLLPPPPALCGASACDVSVRVVVGTEA 480
Db 421 ASSEDSLMTSFLPGPKPGAPFPNGHVGGAGSGLLPPPPALCGASACDVSVRVVVGTEA 480
Qy 481 RVVPGRGICLDLALDLSAPLISQVAPSLFMGSIVQLSQSVTAYMVSAAGLGLVAIFYATQ 540
Db 481 RVVPGRGICLDLALDLSAPLISQVAPSLFMGSIVQLSQSVTAYMVSAAGLGLVAIFYATQ 540
Qy 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

RESULT 4
ID AAY82002 standard; Protein; 553 AA.
XX AAY82002;
XX 13-JUN-2000 (first entry)
XX Human immunogenic prostate tumour protein sequence SEQ ID NO:113.
XX Human; prostate cancer; diagnosis; tumour; gene therapy; detection;
XX immunogenic; cytostatic; vaccine.
XX Homo sapiens.
XX WO2000004149-A2.
XX 27-JAN-2000.
XX 14-JUL-1999; 99WO-US15838.
XX 14-JUL-1998; 98US-0115453.
XX 14-JUL-1998; 98US-0116134.
XX 23-SEP-1998; 98US-0159812.
XX 23-SEP-1998; 98US-0159822.
XX 15-JAN-1999; 99US-0232149.
XX 15-JAN-1999; 99US-0232880.
XX 09-APR-1999; 99US-0288946.
XX (CORI-) CORIXA CORP.
XX Dillon DC, Harlocker SL, Yuguu J, Xu J, Mitcham JL;
XX WPI; 2000-171268/15.
XX New polypeptide useful for treating and diagnosing prostate cancer
XX comprises an immunogenic portion of prostate tumor protein -
XX Claim 3; Page 138-139; 263pp; English.
XX The present invention describes isolated polypeptides, comprising an
XX immunogenic portion of a prostate tumour protein (PTP). The polypeptides
XX and polynucleotides encoding them have cytostatic activity and can be
XX used in vaccines and in gene therapy. The polypeptides and
XX polynucleotides encoding them, antigen presenting cells which express

```

CC the polypeptides, antibodies against the polypeptides and vaccines
 CC comprising them can be used for inhibiting the development of prostate
 CC cancer in a patient. The polypeptides can be used to generate antibodies
 CC or anti-idiotypic antibodies for passive immuno therapy. A portion of
 CC the polynucleotides encoding the polypeptides can be used as a probe or
 CC to modulate the expression of the polypeptides. AAA06241 to AAA06691 and
 CC AA082000 to AA082020 represent sequences used in the exemplification of
 CC the present invention.

XX SQ Sequence 553 AA;
 Query Match 100.0%; Score 2861; DB 21; Length 553;
 Best Local Similarity 100.0%; Pred. No. 2.2e-269;
 Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVQRLWVSRLRRHKAQLLLVNLITFGLEVCLAAGITYVPPLLEVGVEEKFMTVLGIG 60
 Db 1 MVQRLWVSRLRRHKAQLLLVNLITFGLEVCLAAGITYVPPLLEVGVEEKFMTVLGIG 60
 Qy 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
 Db 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
 Qy 121 ELALLILGVGLDFCGQVCFPTLEALLSDLPDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
 Db 121 ELALLILGVGLDFCGQVCFPTLEALLSDLPDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
 Qy 181 IDWTSALAPYLGTQBECLFGLLTLLFTLCTVAATLLVAEEAALGPTPEAGLSAPLSPH 240
 Db 181 IDWTSALAPYLGTQBECLFGLLTLLFTLCTVAATLLVAEEAALGPTPEAGLSAPLSPH 240
 Qy 241 CCPCRLARLAFRNLCALLPRLHQLCCMRPTRLRLFVAELCSWMAALMTFTLYTDFVGEGL 300
 Db 241 CCPCRLARLAFRNLCALLPRLHQLCCMRPTRLRLFVAELCSWMAALMTFTLYTDFVGEGL 300
 Qy 301 YQGVPRAPGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
 Db 301 YQGVPRAPGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
 Qy 361 AFVPAAGATCLSHSVAVVTASALTGTFPSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 Db 361 AFVPAAGATCLSHSVAVVTASALTGTFPSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 Qy 421 ASSEDSLMTSFLPGKPGCAPPNHGVGAGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
 Db 421 ASSEDSLMTSFLPGKPGCAPPNHGVGAGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
 Qy 481 RVVPRGICLDLALDSAFLLSQVAPSLFVMSIIVOLSQSVTAYMVSAAGLVAIYEATQ 540
 Db 481 RVVPRGICLDLALDSAFLLSQVAPSLFVMSIIVOLSQSVTAYMVSAAGLVAIYEATQ 540
 Qy 541 VVFDKSLAKYSA 553
 Db 541 VVFDKSLAKYSA 553

RESULT 5
 AAU69763
 ID AAU69763 standard; Protein; 553 AA.
 XX AAU69763;
 AC AAU69763;
 XX 30-JAN-2002 (first entry)
 DT Human prostate cDNA encoded protein #3.
 XX Human; prostate cancer; cytostatic; immunostimulant; tumour; immunogen.
 XX Homo sapiens.
 OS WO200173032-A2.
 PN 04-OCT-2001.
 XX PD

XX 27-MAR-2001; 2001WO-US09919.
 XX 27-MAR-2000; 2000US-0536857.
 PR 09-MAY-2000; 2000US-0568100.
 PR 12-MAY-2000; 2000US-0570737.
 PR 13-JUN-2000; 2000US-0593793.
 PR 27-JUN-2000; 2000US-0605783.
 PR 10-AUG-2000; 2000US-0636215.
 PR 29-AUG-2000; 2000US-0651236.
 PR 06-SEP-2000; 2000US-0657279.
 PR 02-OCT-2000; 2000US-0679426.
 PR 10-OCT-2000; 2000US-0685166.
 XX (CORI-) CORIXA CORP.
 PA Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
 PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
 PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;
 XX WPI; 2001-639232/73.
 DR N-PSDB; AAS63557.
 XX New human prostate-specific polypeptides and polynucleotides useful for
 PT the diagnosis and treatment of cancer, especially prostate cancer -
 XX Claim 2; Page 269-270; 579pp; English.
 XX The invention relates to isolated prostate-specific
 CC polynucleotides, polypeptides, fusion proteins of the polypeptides,
 CC antibodies raised against the polypeptides (or antigenic epitopes
 CC derived from them) and antigen-presenting cells expressing the
 CC polypeptides. The antibodies are useful for detecting the presence of
 CC cancer, especially prostate cancer. The polypeptides, polynucleotides and
 CC the antigen-presenting cells are useful for stimulating and/or expanding
 CC T cells specific for a tumour protein, and for inhibiting the development
 CC of cancer especially prostate cancer. Compositions comprising an immune
 CC polynucleotide and/or polypeptide are useful for stimulating an immune
 CC response, and for treating cancer. The oligonucleotide is useful for
 CC detecting cancer. The present sequence is a prostate specific
 CC polypeptide of the invention.

XX SQ Sequence 553 AA;
 Query Match 100.0%; Score 2861; DB 22; Length 553;
 Best Local Similarity 100.0%; Pred. No. 2.2e-269;
 Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVQRLWVSRLRRHKAQLLLVNLITFGLEVCLAAGITYVPPLLEVGVEEKFMTVLGIG 60
 Db 1 MVQRLWVSRLRRHKAQLLLVNLITFGLEVCLAAGITYVPPLLEVGVEEKFMTVLGIG 60
 Qy 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
 Db 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
 Qy 121 ELALLILGVGLDFCGQVCFPTLEALLSDLPDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
 Db 121 ELALLILGVGLDFCGQVCFPTLEALLSDLPDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
 Qy 181 IDWTSALAPYLGTQBECLFGLLTLLFTLCTVAATLLVAEEAALGPTPEAGLSAPLSPH 240
 Db 181 IDWTSALAPYLGTQBECLFGLLTLLFTLCTVAATLLVAEEAALGPTPEAGLSAPLSPH 240
 Qy 241 CCPCRLARLAFRNLCALLPRLHQLCCMRPTRLRLFVAELCSWMAALMTFTLYTDFVGEGL 300
 Db 241 CCPCRLARLAFRNLCALLPRLHQLCCMRPTRLRLFVAELCSWMAALMTFTLYTDFVGEGL 300
 Qy 301 YQGVPRAPGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
 Db 301 YQGVPRAPGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
 Qy 361 AFVPAAGATCLSHSVAVVTASALTGTFPSALQILPYTLASLYHREKQVFLPKYRGDTGG 420

```

Db 361 APPVAAGATCLSHSVAVVTASAAALGFTFSAALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Qy 421 ASSEDSMTSFLPGPKPGAPFPNGHVGAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Db 421 ASSEDSMTSFLPGPKPGAPFPNGHVGAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Qy 481 RVVPGRGICLDLAILDLSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAIYFATQ 540
Db 481 RVVPGRGICLDLAILDLSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAIYFATQ 540
Qy 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

RESULT 6
AAU04961
ID AAU04961 standard; Protein; 553 AA.
XX
AC AAU04961;
XX
DT 24-OCT-2001 (first entry)
XX
DE Human prostate tumour protein L1-12.
XX
KW Human; prostate tumour protein; prostate cancer.
XX
OS Homo sapiens.
XX
PN US6262245-B1.
XX
PD 17-JUL-2001.
XX
PE 25-FEB-1998; 98US-0030607.
XX
PR 25-FEB-1997; 97US-0806099.
PR 01-AUG-1997; 97US-0904804.
PR 09-FEB-1998; 98US-0020956.
XX
PA (CORI-) CORIXA CORP.
XX
PI Xu J, Dillon DC;
XX
WPI; 2001-440862/47.
XX
N-PSDB; RAS10108.
XX
Novel polynucleotide encoding polypeptide comprising a portion of
PT prostate tumour protein useful for inhibiting development of prostate
PT cancer or for treating prostate cancer in a patient
XX
XX
Example 1; Column 125-127; 105pp; English.
XX
The sequence is a partial prostate tumour protein, encoded by a prostate
CC tumour specific cDNA. The DNA is useful for inhibiting the development
CC of prostate cancer or for treating prostate cancer in a patient.
XX
SQ Sequence 553 AA;

Query Match 100.0%; Score 2861; DB 22; Length 553;
Best Local Similarity 100.0%; Pred. No. 2.2e-269;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVQRLWVSRLLRHRKAQLLNLTGFLVCLAGITVVPPLLLEVGVEEKFTMTVLIG 60
Db 1 MVQRLWVSRLLRHRKAQLLNLTGFLVCLAGITVVPPLLLEVGVEEKFTMTVLIG 60
Qy 61 PVGLGVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLGLLCPDRPL 120
Db 61 PVGLGVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLGLLCPDRPL 120
Qy 121 ELALLILGVGLLDFCGQVCFPTFLEALLSDLFRDPDHCRQAYSVYAFMISLGGCLGYLLPA 180

```

```

Db 121 ELALLILGVGLLDFCGQVCFPTFLEALLSDLFRDPDHCRQAYSVYAFMISLGGCLGYLLPA 180
Qy 181 IDWTSALAPYLGTOEECLFGLLTLLIFLTCAATLLVAEEAALGFTBPAGLSAPLSPH 240
Db 181 IDWTSALAPYLGTOEECLFGLLTLLIFLTCAATLLVAEEAALGFTBPAGLSAPLSPH 240
Qy 241 CCPCRLARFNIGALLPRHLQCCRPRTLRLRFLVABLCSWMALMTFTLYTDFVGEGL 300
Db 241 CCPCRLARFNIGALLPRHLQCCRPRTLRLRFLVABLCSWMALMTFTLYTDFVGEGL 300
Qy 301 YQGVPRAEPTGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
Db 301 YQGVPRAEPTGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
Qy 361 APVAAGATCLSHSVAVVTASAAALGFTFSAALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APVAAGATCLSHSVAVVTASAAALGFTFSAALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Qy 421 ASSEDSMTSFLPGPKPGAPFPNGHVGAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Db 421 ASSEDSMTSFLPGPKPGAPFPNGHVGAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Qy 481 RVVPGRGICLDLAILDLSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAIYFATQ 540
Db 481 RVVPGRGICLDLAILDLSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAIYFATQ 540
Qy 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

RESULT 7
AAU01117
ID AAU01117 standard; Protein; 553 AA.
XX
AC AAU01117;
XX
DT 04-OCT-2001 (first entry)
XX
DE Human prostate-specific amino acid sequence L1-12.
XX
KW Human; prostate cancer; prostate-specific; diagnosis; vaccine;
KW cytostatic; gene therapy; metastasis.
XX
OS Homo sapiens.
XX
PN WO200151633-A2.
XX
PD 19-JUL-2001.
XX
PF 16-JAN-2001; 2001WO-US01574.
XX
PR 14-JAN-2000; 2000US-0483672.
XX
PA (CORI-) CORIXA CORP.
XX
PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
PI Kalos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW;
PI Wang A, Meagher MJ;
XX
WPI; 2001-425873/45.
XX
New polynucleotide encoding a prostate-specific protein, for
PT diagnosing, monitoring and treating prostate cancer in a patient and
PT for use in vaccines -
XX
PS Claim 2; Page 267-268; 543pp; English.
XX
The present invention describes polynucleotide sequences (I) which encode
CC prostate-specific proteins (II). (I) and (II) have cytostatic activity,
CC and can be used in vaccine production and gene therapy. (I), (II),
CC antibodies to (II), fusion proteins comprising (II), and isolated
CC T cells prepared using (I) or (II) are used treat cancer in a patient.

```

CC (I) and the antibodies are also used in the detection of cancer in a
 CC patient. The cancer that is diagnosed or treated is particularly
 CC prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
 CC (I) can be used for monitoring the progression of cancer in a patient.
 CC (I) and (II) can also be used to improve diagnostic and therapeutic
 CC methods for prostate cancer. They can indicate the level of metastasis
 CC as well as the prostate volume. AAH93357 to AAH93944 and AAH01115 to
 CC AAH01318 represent polynucleotide and amino acid sequences used in the
 CC exemplification of the present invention.

XX
 XX
 XX Sequence 553 AA;
 Query Match 100.0%; Score 2861; DB 22; Length 553;
 Best Local Similarity 100.0%; Pred. No. 2.2e-269;
 Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVQRLWVSRLLRHRKAQLLLVNLITFGLVCLAAAGITYVPPLLEVGVEEKFMTWLGIG 60
 Db 1 MVQRLWVSRLLRHRKAQLLLVNLITFGLVCLAAAGITYVPPLLEVGVEEKFMTWLGIG 60

Qy 61 PVGLVCVPLLGASADHWGRGRRPFIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
 Db 61 PVGLVCVPLLGASADHWGRGRRPFIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120

Qy 121 ELALLILGVGLDFCGQVCFPTLEALLSDLPDPDHCQRAYSVYAFMISLGGCLGYLLPA 180
 Db 121 ELALLILGVGLDFCGQVCFPTLEALLSDLPDPDHCQRAYSVYAFMISLGGCLGYLLPA 180

Qy 181 IDWDTALAPYLGTOECLFGLITLFLTCVAATLLVAEEAALGPTPEAGLSAPSLSPH 240
 Db 181 IDWDTALAPYLGTOECLFGLITLFLTCVAATLLVAEEAALGPTPEAGLSAPSLSPH 240

Qy 241 CCPCRLAFRNIGALLPRLHQLCCRMPTLRLFLVAELCSWMLMTFTLYTDFVGEGL 300
 Db 241 CCPCRLAFRNIGALLPRLHQLCCRMPTLRLFLVAELCSWMLMTFTLYTDFVGEGL 300

Qy 301 YQGVPRAEPTGTEARRHYDEGVRMGSLGLFLOCAISLVFSLVMDRLVORFGTRAVYLASVA 360
 Db 301 YQGVPRAEPTGTEARRHYDEGVRMGSLGLFLOCAISLVFSLVMDRLVORFGTRAVYLASVA 360

Qy 361 APVVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLVHREKQVLPKYRGDTGG 420
 Db 361 APVVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLVHREKQVLPKYRGDTGG 420

Qy 421 ASSDSLMTSFLPGPKPGAPFPNGHVAGSGSLPPPPALCGASACDVSVRVVVGTEA 480
 Db 421 ASSDSLMTSFLPGPKPGAPFPNGHVAGSGSLPPPPALCGASACDVSVRVVVGTEA 480

Qy 541 VFDPKSLAKYSA 553
 Db 541 VFDPKSLAKYSA 553

RESULT 8
 AAG99002
 ID AAG99002 standard; Protein; 553 AA.
 XX
 XX AAG99002;
 AC
 AC
 XX
 XX
 DT 25-SEP-2001 (first entry)
 XX
 XX Human prostate-specific amino acid sequence L1-12/P501S.
 DE
 XX Human; prostate cancer; therapy; diagnosis; cat eye syndrome;
 KW chromosome 22q11.2; prostate-specific protein; chromosome 1;
 KW prostate specific antigen; PSA.
 XX
 XX Homo sapiens.
 OS
 XX

PN WO200134802-A2.
 XX
 XX 17-MAY-2001.
 XX
 XX 09-NOV-2000; 2000WO-US30904.
 XX
 XX 12-NOV-1999; 99US-0439313.
 PR 18-NOV-1999; 99US-0443686.
 XX
 XX (CORI-) CORIXA CORP.
 PA
 XX
 XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
 PI Kalos MD, Retter MW, Stolk JA, Day CH, Skeiky YAW, Wang A;
 XX
 XX WPI; 2001-308785/32.
 XX
 XX Isolated polypeptide comprising at least an immunogenic portion of a
 PT prostate-specific protein, useful in the diagnosis and therapy of
 PT prostate cancer -
 XX
 XX Claim 3; Page 167-168; 32spp; English.
 PS
 XX The present invention describes an isolated polypeptide (P1) comprising
 CC at least an immunogenic portion of a prostate-specific protein, or its
 CC variant. Also described are polynucleotides (N1) encoding (P1). (P1) and
 CC (N1) have cytostatic activity and can be used in vaccine production.
 CC The polypeptides, nucleic acids and antibodies from the present
 CC invention are useful in the diagnosis and therapy of prostate cancer.
 CC Prostate specific genes P704P, P712P, P774P, P775P and B305D are located
 CC in a genomic region on chromosome 22q11.2 known as the Cat Eye Syndrome
 CC region. Prostate specific antigen (PSA) P501S was located on
 CC chromosome 1. AAH84671 to AAH85143 and AAG99000 to AAG99077 represent
 CC of the present invention.
 XX
 XX Sequence 553 AA;
 Query Match 100.0%; Score 2861; DB 22; Length 553;
 Best Local Similarity 100.0%; Pred. No. 2.2e-269;
 Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVQRLWVSRLLRHRKAQLLLVNLITFGLVCLAAAGITYVPPLLEVGVEEKFMTWLGIG 60
 Db 1 MVQRLWVSRLLRHRKAQLLLVNLITFGLVCLAAAGITYVPPLLEVGVEEKFMTWLGIG 60

Qy 61 PVGLVCVPLLGASADHWGRGRRPFIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
 Db 61 PVGLVCVPLLGASADHWGRGRRPFIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120

Qy 121 ELALLILGVGLDFCGQVCFPTLEALLSDLPDPDHCQRAYSVYAFMISLGGCLGYLLPA 180
 Db 121 ELALLILGVGLDFCGQVCFPTLEALLSDLPDPDHCQRAYSVYAFMISLGGCLGYLLPA 180

Qy 181 IDWDTALAPYLGTOECLFGLITLFLTCVAATLLVAEEAALGPTPEAGLSAPSLSPH 240
 Db 181 IDWDTALAPYLGTOECLFGLITLFLTCVAATLLVAEEAALGPTPEAGLSAPSLSPH 240

Qy 241 CCPCRLAFRNIGALLPRLHQLCCRMPTLRLFLVAELCSWMLMTFTLYTDFVGEGL 300
 Db 241 CCPCRLAFRNIGALLPRLHQLCCRMPTLRLFLVAELCSWMLMTFTLYTDFVGEGL 300

Qy 301 YQGVPRAEPTGTEARRHYDEGVRMGSLGLFLOCAISLVFSLVMDRLVORFGTRAVYLASVA 360
 Db 301 YQGVPRAEPTGTEARRHYDEGVRMGSLGLFLOCAISLVFSLVMDRLVORFGTRAVYLASVA 360

Qy 361 APVVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLVHREKQVLPKYRGDTGG 420
 Db 361 APVVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLVHREKQVLPKYRGDTGG 420

Qy 421 ASSDSLMTSFLPGPKPGAPFPNGHVAGSGSLPPPPALCGASACDVSVRVVVGTEA 480
 Db 421 ASSDSLMTSFLPGPKPGAPFPNGHVAGSGSLPPPPALCGASACDVSVRVVVGTEA 480

QY 481 RVVPRGICLDLAILDSAFLLSQVAPSLFMGSIIVLSQSVTAYVMVSAAGLGLVIAIYFATQ 540
DB 481 RVVPRGICLDLAILDSAFLLSQVAPSLFMGSIIVLSQSVTAYVMVSAAGLGLVIAIYFATQ 540
QY 541 VVFDKSLAKYSA 553
DB 541 VVFDKSLAKYSA 553

RESULT 9
ID AAG62150 standard; Protein; 553 AA.
XX AAG62150;
AC AAG62150;
DT 06-JUL-2001 (first entry)
XX Human P501S inventive antigen SEQ ID NO: 333.
XX Human; mouse; immunotherapy; cancer; leukaemia; Wt1; Wilm's tumour gene;
KW chromosome 11p13; zinc finger transcription factor.
XX Homo sapiens.
XX WO200125273-A2.
XX 12-APR-2001.
XX 04-OCT-2000; 2000WO-US27465.
XX 04-OCT-1999; 99US-0157459.
XX (CORI-) CORIXA CORP.
XX Skeiky YAW, Xu J, Cheever MA, Reed SG;
XX WPI; 2001-328324/34.
XX Polypeptide comprising part of the Wilm's Tumour gene product sequence is
PT used in the diagnosis and treatment of malignant diseases e.g. leukemia
PT and cancer associated with Wt1 -
XX Disclosure; Page 212-213; 228pp; English.
XX The present invention describes compositions comprising peptides derived
CC from the Wilm's tumour protein Wt1 and methods for their use in treating
CC malignant diseases. Peptides derived from both the murine and human Wt1
CC proteins are provided. The human Wt1 gene is found on chromosome 11p13,
CC and the protein was shown to be a zinc finger transcription factor. The
CC immunogenic peptides of the invention are particularly useful in the
CC diagnosis and treatment of cancer and leukaemia. The present sequence is
CC a polypeptide described in the exemplification of the invention.

QY Sequence 553 AA;
SQ

Query Match 100.0%; Score 2861; DB 22; Length 553;
Best Local Similarity 100.0%; Pred. No. 2.2e-269;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWSRLRHRKAQLLVNLTLTFLGVLVCLAAAGITVPPLLLEVGVEEKFTMTVLGIG 60
DB 1 MVQRLWSRLRHRKAQLLVNLTLTFLGVLVCLAAAGITVPPLLLEVGVEEKFTMTVLGIG 60

QY 61 PVGLVLCVPLIGSADHWGRYGRRRPPIWALSIGILLSLFLIPRAGWLAGLLCPDRPL 120
DB 61 PVGLVLCVPLIGSADHWGRYGRRRPPIWALSIGILLSLFLIPRAGWLAGLLCPDRPL 120

QY 121 ELALLILGVGLDFCGQVCFPLEALLSDLFDPDHCRCQAVSVYAFWISLGGCLGYLLPA 180
DB 121 ELALLILGVGLDFCGQVCFPLEALLSDLFDPDHCRCQAVSVYAFWISLGGCLGYLLPA 180

QY 181 IDWTSALAPYLGTQECLFGLLTLLIFLTCVAATLLVAEEAALGPTPAEGLSAPSLSPH 240

DB 181 IDWTSALAPYLGTQECLFGLLTLLIFLTCVAATLLVAEEAALGPTPAEGLSAPSLSPH 240
QY 241 CCPCRLARLAFRNLGALLPRLHQLCCMRPTRLRLRFVAELCSWMALMTFTFTYTDVGBGL 300
DB 241 CCPCRLARLAFRNLGALLPRLHQLCCMRPTRLRLRFVAELCSWMALMTFTFTYTDVGBGL 300
QY 301 YQGVPRAPGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
DB 301 YQGVPRAPGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
QY 361 AFPVAAGATCLSHSVAVVTASALTGTFTSALQTLPTTLASLYHREKOVFLPKYRGDTGG 420
DB 361 AFPVAAGATCLSHSVAVVTASALTGTFTSALQTLPTTLASLYHREKOVFLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGAGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
DB 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGAGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
QY 481 RVVPRGICLDLAILDSAFLLSQVAPSLFMGSIIVLSQSVTAYVMVSAAGLGLVIAIYFATQ 540
DB 481 RVVPRGICLDLAILDSAFLLSQVAPSLFMGSIIVLSQSVTAYVMVSAAGLGLVIAIYFATQ 540
QY 541 VVFDKSLAKYSA 553
DB 541 VVFDKSLAKYSA 553

RESULT 10
ID AAB74800 standard; Protein; 553 AA.
XX AAB74800;
AC AAB74800;
DT 14-JUN-2001 (first entry)
XX Prostate tumour antigen predicted amino acid sequence for LI-12.
XX Human; prostate tumour antigen; prostate tumour; therapy; diagnosis;
KW prostate cancer; immunogenic; cytostatic; vaccine.
XX Homo sapiens.
XX WO200125272-A2.
XX 12-APR-2001.
XX 04-OCT-2000; 2000WO-US27464.
XX 04-OCT-1999; 99US-0157455.
XX (CORI-) CORIXA CORP.
XX Xu J, Skeiky YAW, Reed SG, Cheever MA;
XX WPI; 2001-245062/25.
XX N-PSDB; AAH02530.
XX Prostate specific protein and its encoding polynucleotide, useful for
PT the treatment and diagnosis of prostate cancer -
XX Claim 3; Page 157-158; 276pp; English.
XX The present invention describes an isolated polypeptide (I) comprising
CC at least an immunogenic portion of a prostate tumour antigen protein or
CC its variant. (I) have cytostatic activity and can be used in vaccine
CC production. (II), prostate tumour antigen polynucleotides, an antigen
CC presenting cell (APC e.g. a dendritic cell) that expresses (I), and a
CC pharmaceutical composition containing (I) are useful for inhibiting the
CC development of cancer in a patient. Antibodies specific for prostate
CC specific proteins and oligonucleotides that hybridise to a
CC polynucleotide that encodes a prostate specific protein are useful
CC for detecting the presence or absence of a cancer or monitoring the
CC progression the progression of a cancer, especially prostate cancer.

CC AAH02422 to AAH2872, AAB74798 to AAB74821 and AAB74830 are sequences
CC used in the exemplification of the present invention.

XX
SQ Sequence 553 AA;

Query Match 100.0%; Score 2861; DB 22; Length 553;
Best Local Similarity 100.0%; Pred. No. 2.2e-269;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWSRLRRHKAQQLLVNLTGLEVCLAAGITYVPPLELVGVVEEKFTMWLGIG 60
DB 1 MVQRLWSRLRRHKAQQLLVNLTGLEVCLAAGITYVPPLELVGVVEEKFTMWLGIG 60

QY 61 PVLGLVCPVLLGSASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDRPL 120
DB 61 PVLGLVCPVLLGSASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDRPL 120

QY 121 ELALLILGVGLDFCGQVCFPLEALLSDLPDPDHCQAYSVVAFMISLGGCLGYLLPA 180
DB 121 ELALLILGVGLDFCGQVCFPLEALLSDLPDPDHCQAYSVVAFMISLGGCLGYLLPA 180

QY 181 IDWTSALAPYLGTQBECLFGLLTLLIFLTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
DB 181 IDWTSALAPYLGTQBECLFGLLTLLIFLTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240

QY 241 CCPCARLAFRNLGALLPRHLQCCMRPTRLRLFVAELCSWMALMTFTLYTDFVGEGL 300
DB 241 CCPCARLAFRNLGALLPRHLQCCMRPTRLRLFVAELCSWMALMTFTLYTDFVGEGL 300

QY 301 YQGVPRAEPTGTEARRHYDEGVNMGSLGLFLOCAISLVFSLVMDRLVORFGTRAVYLASVA 360
DB 301 YQGVPRAEPTGTEARRHYDEGVNMGSLGLFLOCAISLVFSLVMDRLVORFGTRAVYLASVA 360

QY 361 AFPVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLYHREKQVLPKYRGDTGG 420
DB 361 AFPVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLYHREKQVLPKYRGDTGG 420

QY 421 ASSEDSLMTSFLPGPKPGAPFNGHVAGGSGLLPPPPALCGASACDVSVRVVVGTEA 480
DB 421 ASSEDSLMTSFLPGPKPGAPFNGHVAGGSGLLPPPPALCGASACDVSVRVVVGTEA 480

QY 481 RVVPGRGICLDLALDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLVAIYFATQ 540
DB 481 RVVPGRGICLDLALDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLVAIYFATQ 540

RESULT 11
ABG94411
ID ABG94411 standard; Protein; 553 AA.

XX AC ABG94411;

XX DT 27-NOV-2002 (first entry)

XX DE Human prostate tumour protein partial sequence #3.

XX KW Human; immunogenic; prostate protein; prostate tumour protein;
XX KW prostate cancer; cytostatic; vaccine.

XX OS Homo sapiens.

XX EN US2002090372-A1.

XX PD 11-JUL-2002.

XX PF 14-JUL-1998; 98US-0115453.

XX PR 25-FEB-1997; 97US-0806099.

XX PR 01-AUG-1997; 97US-0904804.

PR 09-FEB-1998; 98US-0020956.
PR 25-FEB-1998; 98US-0030607.

XX (XUJ/) XU J.
XX (DILL/) DILLON D C.

XX Xu J, Dillon DC;
XX WPI; 2002-642373/69.
XX N-PSDB; ABS21254.

Novel polypeptides useful as vaccines for inhibiting prostate cancer development, comprise an immunogenic portion of prostate protein -

Example 1; Page 59-61; 101pp; English.

The present invention relates to a new polypeptide comprising an immunogenic portion of a prostate protein. The invention is useful for inhibiting the development of prostate cancer in a patient. The invention is also useful as markers for diagnosing prostate cancer and for monitoring disease progression in patients. The present amino acid sequence represents a human prostate tumour protein.

SQ Sequence 553 AA;

Query Match 100.0%; Score 2861; DB 23; Length 553;
Best Local Similarity 100.0%; Pred. No. 2.2e-269;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWSRLRRHKAQQLLVNLTGLEVCLAAGITYVPPLELVGVVEEKFTMWLGIG 60
DB 1 MVQRLWSRLRRHKAQQLLVNLTGLEVCLAAGITYVPPLELVGVVEEKFTMWLGIG 60

QY 61 PVLGLVCPVLLGSASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDRPL 120
DB 61 PVLGLVCPVLLGSASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDRPL 120

QY 121 ELALLILGVGLDFCGQVCFPLEALLSDLPDPDHCQAYSVVAFMISLGGCLGYLLPA 180
DB 121 ELALLILGVGLDFCGQVCFPLEALLSDLPDPDHCQAYSVVAFMISLGGCLGYLLPA 180

QY 181 IDWTSALAPYLGTQBECLFGLLTLLIFLTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
DB 181 IDWTSALAPYLGTQBECLFGLLTLLIFLTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240

QY 241 CCPCARLAFRNLGALLPRHLQCCMRPTRLRLFVAELCSWMALMTFTLYTDFVGEGL 300
DB 241 CCPCARLAFRNLGALLPRHLQCCMRPTRLRLFVAELCSWMALMTFTLYTDFVGEGL 300

QY 301 YQGVPRAEPTGTEARRHYDEGVNMGSLGLFLOCAISLVFSLVMDRLVORFGTRAVYLASVA 360
DB 301 YQGVPRAEPTGTEARRHYDEGVNMGSLGLFLOCAISLVFSLVMDRLVORFGTRAVYLASVA 360

QY 361 AFPVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLYHREKQVLPKYRGDTGG 420
DB 361 AFPVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLYHREKQVLPKYRGDTGG 420

QY 421 ASSEDSLMTSFLPGPKPGAPFNGHVAGGSGLLPPPPALCGASACDVSVRVVVGTEA 480
DB 421 ASSEDSLMTSFLPGPKPGAPFNGHVAGGSGLLPPPPALCGASACDVSVRVVVGTEA 480

QY 481 RVVPGRGICLDLALDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLVAIYFATQ 540
DB 481 RVVPGRGICLDLALDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLVAIYFATQ 540

QY 541 VFEDKSLAKYSA 553
DB 541 VFEDKSLAKYSA 553

RESULT 12
ABG76665
ID ABG76665 standard; Protein; 553 AA.

XX ABG76665;
 AC 05-NOV-2002 (first entry)
 DT Prostate tumour protein #3.
 DE Human; prostate tumour; immunotherapy; prostate cancer.
 KW Homo sapiens.
 XX OS US2002081580-A1.
 XX PN 27-JUN-2002.
 XX PD 25-FEB-1998; 98US-0030606.
 XX PF 25-FEB-1997; 97US-0806596.
 PR 01-AUG-1997; 97US-0904809.
 PR 09-FEB-1998; 98US-0020747.
 XX (XUJJ/) XU J.
 PA (DILL/) DILLON D C.
 XX PI Xu J, Dillon DC;
 XX DR WPI; 2002-607662/65.
 XX DT Detecting prostate cancer comprises contacting a sample with an agent
 PT capable of binding to a polypeptide with an immunogenic portion of a
 PT prostate protein, oligonucleotide primers or a probe specific for DNA
 PT encoding the polypeptide
 XX Example 1; Page 65-67; 11pp; English.
 XX The invention relates to a method of detecting prostate cancer by
 CC contacting a biological sample from a patient with: (a) a binding agent
 CC that binds to a polypeptide having an immunogenic portion of a prostate
 CC protein or its variant; (b) 2 oligonucleotide primers, where 1 of the
 CC oligonucleotides is specific for a DNA encoding the polypeptide of (a);
 CC or (c) an oligonucleotide probe specific for a DNA molecule encoding the
 CC polypeptide of (a). The method and polypeptides are useful for
 CC diagnosing, treating, particularly by immunotherapy, monitoring the
 CC progression, and inhibiting the development of prostate cancer in a
 CC patient. The polypeptides may be used to generate antibodies useful for
 CC the diagnosis and monitoring of prostate cancer. ABG76663-ABG76669
 CC represent human prostate tumour protein sequences of the invention.
 XX SQ Sequence 553 AA;
 Query Match 100.0%; Score 2861; DB 23; Length 553;
 Best Local Similarity 100.0%; Pred. No. 2.2e-269; Mismatches 0; Gaps 0;
 Matches 553; Conservative 0; Indels 0; Gaps 0;
 QY 1 MVQRLWSRLRHRKAQLLVNLTLFGLVCLAAAGITVVPPLLEVGVEEKFTMWLGIG 60
 DB 1 MVQRLWSRLRHRKAQLLVNLTLFGLVCLAAAGITVVPPLLEVGVEEKFTMWLGIG 60
 QY 61 PVLGLVCPVLLGSASDHWGRYGRRRPFIWALSGLILLSLFLIPRAGWLAGLLCPDRPL 120
 DB 61 PVLGLVCPVLLGSASDHWGRYGRRRPFIWALSGLILLSLFLIPRAGWLAGLLCPDRPL 120
 QY 121 ELALLIIGVGLDRCQVCFPTLEALLSDLPDRDHCQAVSVVAFMISLGGCLGYLLPA 180
 DB 121 ELALLIIGVGLDRCQVCFPTLEALLSDLPDRDHCQAVSVVAFMISLGGCLGYLLPA 180
 QY 181 IDWTSALAPYLQEQECLFGLLTILFUTCAATLLVAEEAALGPTPEAGLSAPSLSPH 240
 DB 181 IDWTSALAPYLQEQECLFGLLTILFUTCAATLLVAEEAALGPTPEAGLSAPSLSPH 240
 QY 241 CCPCRLAFRNLCALLPRLHQLCCMRPTRLRLRFVaelCSWMAIMFTFLFYTFDFVGEGL 300
 DB 241 CCPCRLAFRNLCALLPRLHQLCCMRPTRLRLRFVaelCSWMAIMFTFLFYTFDFVGEGL 300

QY 301 YQGVPRAPGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVQRFQFTRAVYLASVA 360
 DB 301 YQGVPRAPGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVQRFQFTRAVYLASVA 360
 QY 361 APVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 DB 361 APVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 QY 421 ASSEDSLMTSFLPGKPGAPPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPEPTA 480
 DB 421 ASSEDSLMTSFLPGKPGAPPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGPEPTA 480
 QY 481 RVPEGRGICLDLAIDLDAFLLSQVAPSLFMGSIYVQLSQSVTAYMVVSAAGLGLVAIFYATQ 540
 DB 481 RVPEGRGICLDLAIDLDAFLLSQVAPSLFMGSIYVQLSQSVTAYMVVSAAGLGLVAIFYATQ 540
 QY 541 VVFDKSDIAKYSA 553
 DB 541 VVFDKSDIAKYSA 553

RESULT 13
 ABB77575
 ID ABB77575 standard; Protein; 553 AA.
 XX AC ABB77575;
 XX DT 30-AUG-2002 (first entry)
 XX DE Human mast cell related splice variant protein MC14 SEQ ID NO 13.
 XX KW Human; mast cell; MC; antiallergic; antiinflammatory; antiasthmatic;
 XX KV vasotropic; dermatological; allergy; hypersensitivity; rhinitis; asthma.
 XX OS Homo sapiens.
 XX PN WO200246389-A2.
 XX PD 13-JUN-2002.
 XX PF 07-DEC-2001; 2001WO-US46180.
 XX PR 08-DEC-2000; 2000US-251835P.
 XX PR 14-MAR-2001; 2001US-275479P.
 XX PR 28-MAR-2001; 2001US-279115P.
 XX PR 02-APR-2001; 2001US-280143P.
 XX PA (UNIO) UCB SA.
 XX PI Nocka K, Pirozzi G, Einstein R;
 XX WPI; 2002-508560/54.
 DR N-PSDB; ABN81324.
 XX Novel isolated nucleic acids that are differentially expressed in mast
 PT calls in patients with allergic hypersensitivity, encoding proteins
 PT associated with mast cell regranulation and allergic hypersensitivity
 XX Claim 31; Page 117-119; 119pp; English.
 XX The invention relates to isolated nucleic acid (ABN81319-ABN81324),
 CC corresponding to genes differentially expressed in mast cells following
 CC activation or in patients with allergic hypersensitivity disease, (I)
 CC that encodes proteins (ABN7569-ABN7575) (II) or a protein fragment of
 CC (II) if at least 6 amino acids. (II) is useful for identifying binding
 CC partners. (I) or (II) is useful for diagnosing or treating a disease
 CC state (e.g. allergic hypersensitivity, seasonal rhinitis, asthma,
 CC urticaria or atopic dermatitis or mastocytosis) in a subject which
 CC involves determining the level of expression of (I) or (II). A computer
 CC system, comprising a database containing information identifying the
 CC expression level in a tissue or at least one mast cell of (I), is useful

CC for presenting information to identify the relative expression level of
 CC (I). (II) is used as a marker to detect, diagnose or identify an allergic
 CC response in a patient. The protein can also serve as a target that
 CC modulate gene expression or activity and as an antigen to raise
 CC polyclonal or monoclonal antibodies. (II) is useful for identifying
 CC agents that modulate expression of the protein or agents, such as
 CC agonists or antagonists. The agonists or antagonists are useful for
 CC modulating biological activity and function of (II) and thus are useful
 CC for alleviating disease conditions such as allergic hypersensitivity,
 CC seasonal rhinitis, asthma, urticaria, atopic dermatitis or mastocytosis.
 XX
 SQ Sequence 553 AA;

Query Match 100.0%; Score 2861; DB 23; Length 553;
 Best Local Similarity 100.0%; Pred. No. 2.2e-269;
 Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLRRHKAQLLVNLITFGLEVCIAAGITVPPLLEVGVEEKFMTMVLIG 60
 DB 1 MVQRLWVSRLRRHKAQLLVNLITFGLEVCIAAGITVPPLLEVGVEEKFMTMVLIG 60
 QY 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDRPL 120
 DB 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDRPL 120
 QY 121 ELALLILGVGLDFCGQVCFPTLEALLSDLPDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
 DB 121 ELALLILGVGLDFCGQVCFPTLEALLSDLPDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
 QY 181 IDWTSALAPYLGTQECLEFGLLTFLITCVAATLLVAEEAALGPTEPAEGLSAPLSPH 240
 DB 181 IDWTSALAPYLGTQECLEFGLLTFLITCVAATLLVAEEAALGPTEPAEGLSAPLSPH 240
 QY 241 CCPCRLAFRNIGALLPRHLQCCMRPTRLRLFLVAELCSWMAIMTFTLYTDFVGEGL 300
 DB 241 CCPCRLAFRNIGALLPRHLQCCMRPTRLRLFLVAELCSWMAIMTFTLYTDFVGEGL 300
 QY 301 YQGVPRAPGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVQRFTRAVYLASVA 360
 DB 301 YQGVPRAPGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVQRFTRAVYLASVA 360
 QY 361 APVAAGATCLSHSVAVVTAASALTGTFTSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 DB 361 APVAAGATCLSHSVAVVTAASALTGTFTSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
 QY 421 ASSEDSLMTSPLPGPKCAPPPNGHVAGSGSGLLPPPPALCGASACDVSVRVVVGTEA 480
 DB 421 ASSEDSLMTSPLPGPKCAPPPNGHVAGSGSGLLPPPPALCGASACDVSVRVVVGTEA 480
 QY 481 RVVPRGICLDLAILDSAFLLSQVAPSLFMGSIYQLSQSVTAYMVSAAGLGLVAIYATQ 540
 DB 481 RVVPRGICLDLAILDSAFLLSQVAPSLFMGSIYQLSQSVTAYMVSAAGLGLVAIYATQ 540
 QY 541 WFDKSLAKYSA 553
 DB 541 WFDKSLAKYSA 553

RESULT 14
 ABG61900
 ID ABG61900 standard; Protein; 553 AA.
 XX
 AC ABG61900;
 XX
 DT 15-AUG-2002 (first entry)
 XX
 DE Prostate cancer-associated protein #101.
 XX
 KW Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.
 XX
 OS Mammalia.
 XX
 PN WO20020268-A2.

XX 18-APR-2002.
 PD 12-OCT-2001; 2001WO-US32045.
 XX
 PF 13-OCT-2000; 2000US-0687576.
 XX
 PR 08-DEC-2000; 2000US-0733288.
 PR 08-DEC-2000; 2000US-0733742.
 PR 24-JAN-2001; 2001US-263957P.
 PR 16-MAR-2001; 2001US-276791P.
 PR 06-APR-2001; 2001US-281922P.
 PR 24-APR-2001; 2001US-286214P.
 PR 30-APR-2001; 2001US-0847046.
 PR 04-MAY-2001; 2001US-288589P.
 XX
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
 XX
 WPI; 2002-471335/50.
 DR N-PSDB; ABK92217.
 XX
 PT Detecting a prostate cancer-associated transcript in a cell in a
 PT patient, useful for diagnosing prostate cancer (PC) or screening
 PT modulators of PC, by determining if prostate cancer-associated genes
 PT are expressed in a prostate tissue -
 XX
 PS Claim 27; Page 386; 436pp; English.
 XX
 CC The present invention relates to methods of detecting a prostate
 CC cancer-associated transcript in a cell from a patient. The method
 CC comprises contacting a biological sample from the patient with
 CC prostate cancer-associated polynucleotides (designated PC genes) that
 CC selectively hybridize to a sequence that is at least 80% identical
 CC to them. The prostate cancer-associated polynucleotide sequences
 CC are differentially expressed in prostate tumour tissue or in
 CC prostate cancer and are derived from the tissues of various
 CC organisms such as humans or other mammals (e.g. mice, sheep and dogs).
 CC The methods of the invention are useful for diagnosing and treating
 CC prostate cancer in mammals. The prostate cancer-associated genes are
 CC useful for diagnosing or treating prostate cancer, as well as for
 CC identifying modulators of prostate cancer or agents that inhibit
 CC prostate cancer. The nucleic acid sequences are particularly useful
 CC in gene therapy, as a vaccine or in antisense applications.
 CC ABG61800-ABG61944 represent prostate cancer-associated proteins.
 XX
 SQ Sequence 553 AA;

Query Match 100.0%; Score 2861; DB 23; Length 553;
 Best Local Similarity 100.0%; Pred. No. 2.2e-269;
 Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MVQRLWVSRLRRHKAQLLVNLITFGLEVCIAAGITVPPLLEVGVEEKFMTMVLIG 60
 DB 1 MVQRLWVSRLRRHKAQLLVNLITFGLEVCIAAGITVPPLLEVGVEEKFMTMVLIG 60
 QY 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDRPL 120
 DB 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDRPL 120
 QY 121 ELALLILGVGLDFCGQVCFPTLEALLSDLPDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
 DB 121 ELALLILGVGLDFCGQVCFPTLEALLSDLPDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
 QY 181 IDWTSALAPYLGTQECLEFGLLTFLITCVAATLLVAEEAALGPTEPAEGLSAPLSPH 240
 DB 181 IDWTSALAPYLGTQECLEFGLLTFLITCVAATLLVAEEAALGPTEPAEGLSAPLSPH 240
 QY 241 CCPCRLAFRNIGALLPRHLQCCMRPTRLRLFLVAELCSWMAIMTFTLYTDFVGEGL 300
 DB 241 CCPCRLAFRNIGALLPRHLQCCMRPTRLRLFLVAELCSWMAIMTFTLYTDFVGEGL 300

Search completed: January 25, 2004, 17:03:48
Job time : 50 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2004, 17:02:52 ; Search time 21 Seconds
(without alignments)
1114.186 Million cell updates/sec

Title: US-09-593-793A-113
Perfect score: 2861
Sequence: 1 MVQRLWVSRLRHRKAQLL.....AIYFATQVVFQKSLAKYSA 553

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2861	100.0	553	US-09-020-956-113	Sequence 113, App
2	2861	100.0	553	US-09-030-607-113	Sequence 113, App
3	2861	100.0	553	US-09-439-313-113	Sequence 113, App
4	2861	100.0	553	US-09-352-616A-113	Sequence 113, App
5	2861	100.0	553	US-09-602-877A-101	Sequence 101, App
6	2861	100.0	553	US-09-232-149A-113	Sequence 113, App
7	1287	45.0	255	US-09-071-710-36	Sequence 36, Appl
8	1287	45.0	255	US-09-525-397-36	Sequence 36, Appl
9	452	15.8	84	US-09-439-313-571	Sequence 571, Appl
10	330	11.5	516	US-08-356-340-4	Sequence 4, Appl
11	330	11.5	516	US-08-786-555-4	Sequence 4, Appl
12	324.5	11.3	525	US-08-356-340-2	Sequence 2, Appl
13	324.5	11.3	525	US-08-786-555-2	Sequence 2, Appl
14	312	10.9	58	US-09-439-313-547	Sequence 547, App
15	304	10.6	56	US-09-439-313-564	Sequence 564, App
16	243	8.5	44	US-09-071-710-37	Sequence 37, Appl
17	243	8.5	44	US-09-525-397-37	Sequence 37, Appl
18	157	5.5	29	US-09-439-313-546	Sequence 546, App
19	146	5.1	27	US-09-071-710-39	Sequence 39, Appl
20	146	5.1	27	US-09-525-397-39	Sequence 39, Appl
21	146	5.1	27	US-09-439-313-566	Sequence 566, App
22	137	4.8	25	US-09-439-313-520	Sequence 520, App
23	135	4.7	27	US-09-071-710-38	Sequence 38, Appl
24	135	4.7	27	US-09-525-397-38	Sequence 38, Appl
25	133	4.6	409	US-09-252-991A-30461	Sequence 30461, A
26	126	4.4	432	US-09-328-352-4131	Sequence 4131, Ap
27	124.5	4.4	444	US-09-252-991A-20446	Sequence 20446, A

ALIGNMENTS

RESULT 1

US-09-020-956-113
; Sequence 113, Application US/09020956
; Patent No. 6261562
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO
; NUMBER OF SEQUENCES: 178
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/020,956
; FILING DATE: 09-FEB-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.427C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 113:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 553 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; US-09-020-956-113

Query Match 100.0%; Score 2861; DB 3; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.1e-270;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 MVQRLWVSRLRHRKAQLLVLITLFCLEVCAGIYVPPLLLEVGVEEKFMTVLIG 60
Db 1 MVQRLWVSRLRHRKAQLLVLITLFCLEVCAGIYVPPLLLEVGVEEKFMTVLIG 60

QY 61 PVGLVCPVLGSSADHWGRYGRRRPFIFWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
DB 61 PVGLVCPVLGSSADHWGRYGRRRPFIFWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
QY 121 ELALLILGVLLDFCGQVCFPLEALLSDIFRDPDHCRCQAYSVVAFMISLGGCLGYLLPA 180
DB 121 ELALLILGVLLDFCGQVCFPLEALLSDIFRDPDHCRCQAYSVVAFMISLGGCLGYLLPA 180
QY 181 IDWDTALAPYLGTOECLFGLTLIFLTCVAATLLVAEEAALGPTPEAEGLSAPSLSPH 240
DB 181 IDWDTALAPYLGTOECLFGLTLIFLTCVAATLLVAEEAALGPTPEAEGLSAPSLSPH 240
QY 241 CCPCRLAFRNLCALLPRHLQCCMRPTRLRLFVAELCSWMLMTFTFLFYTDVFGEGL 300
DB 241 CCPCRLAFRNLCALLPRHLQCCMRPTRLRLFVAELCSWMLMTFTFLFYTDVFGEGL 300
QY 301 YQGVPRAEPTGEARRHYDEGVRMGSLGLFLOCAISLVFSLVMDRLVORFGTRAVYLASVA 360
DB 301 YQGVPRAEPTGEARRHYDEGVRMGSLGLFLOCAISLVFSLVMDRLVORFGTRAVYLASVA 360
QY 361 APFVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
DB 361 APFVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGPKGAPPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTA 480
DB 421 ASSEDSLMTSFLPGPKGAPPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTA 480
QY 481 RVVPRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAGLGLVAIYATQ 540
DB 481 RVVPRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAGLGLVAIYATQ 540
QY 541 VVFDKSDLAKYSA 553
DB 541 VVFDKSDLAKYSA 553

RESULT 2

US-09-030-607-113
; Sequence 113, Application US/09030607
; Patent No. 6262245
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO
; NUMBER OF SEQUENCES: 224
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/030,607
; FILING DATE: 25-FEB-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.427C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 113:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 553 amino acids

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
US-09-030-607-113

Query Match 100.0%; Score 2861; DB 3; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.1e-270;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVORLWVSRLLRRKKAQLLLVNLITFGLVCLAAAGTYVPPLLEVEGVBEKFTMWLGIG 60
DB 1 MVORLWVSRLLRRKKAQLLLVNLITFGLVCLAAAGTYVPPLLEVEGVBEKFTMWLGIG 60
QY 61 PVGLVCPVLGSSADHWGRYGRRRPFIFWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
DB 61 PVGLVCPVLGSSADHWGRYGRRRPFIFWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
QY 121 ELALLILGVLLDFCGQVCFPLEALLSDIFRDPDHCRCQAYSVVAFMISLGGCLGYLLPA 180
DB 121 ELALLILGVLLDFCGQVCFPLEALLSDIFRDPDHCRCQAYSVVAFMISLGGCLGYLLPA 180
QY 181 IDWDTALAPYLGTOECLFGLTLIFLTCVAATLLVAEEAALGPTPEAEGLSAPSLSPH 240
DB 181 IDWDTALAPYLGTOECLFGLTLIFLTCVAATLLVAEEAALGPTPEAEGLSAPSLSPH 240
QY 241 CCPCRLAFRNLCALLPRHLQCCMRPTRLRLFVAELCSWMLMTFTFLFYTDVFGEGL 300
DB 241 CCPCRLAFRNLCALLPRHLQCCMRPTRLRLFVAELCSWMLMTFTFLFYTDVFGEGL 300
QY 301 YQGVPRAEPTGEARRHYDEGVRMGSLGLFLOCAISLVFSLVMDRLVORFGTRAVYLASVA 360
DB 301 YQGVPRAEPTGEARRHYDEGVRMGSLGLFLOCAISLVFSLVMDRLVORFGTRAVYLASVA 360
QY 361 APFVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
DB 361 APFVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGPKGAPPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTA 480
DB 421 ASSEDSLMTSFLPGPKGAPPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTA 480
QY 481 RVVPRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAGLGLVAIYATQ 540
DB 481 RVVPRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAGLGLVAIYATQ 540
QY 541 VVFDKSDLAKYSA 553
DB 541 VVFDKSDLAKYSA 553

RESULT 3

US-09-439-313-113
; Sequence 113, Application US/09439313
; Patent No. 6329505
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yuqui
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Solk, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.427C9
; CURRENT APPLICATION NUMBER: US/09/439,313

```

; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-439-313-113

Query Match      100.0%; Score 2861; DB 4; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.1e-270;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVQRLWVSRLRHRKQAQLLLVNLTFGLVCLAAAGITVPPLLELVGVVEEKFMVTLGIG 60
Db 1 MVQRLWVSRLRHRKQAQLLLVNLTFGLVCLAAAGITVPPLLELVGVVEEKFMVTLGIG 60

Qy 61 PVGLVCVPLIGSADHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
Db 61 PVGLVCVPLIGSADHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120

Qy 121 ELALLILGVGLDFCGQVCFPLEALLSDLPDPDCHQAYSVYAFMISLGGCLGYLLPA 180
Db 121 ELALLILGVGLDFCGQVCFPLEALLSDLPDPDCHQAYSVYAFMISLGGCLGYLLPA 180

Qy 181 IDWTSALAPYLGTQEECLFGLLTLLIFLTCAATLLVAEEAALGPTPEAGLSAPSLSPH 240
Db 181 IDWTSALAPYLGTQEECLFGLLTLLIFLTCAATLLVAEEAALGPTPEAGLSAPSLSPH 240

Qy 241 CCPCRLAFRNIGALLPRLHQLCCMRPTRLRLFVAELCSWMLMTFTLYTDFVGEGL 300
Db 241 CCPCRLAFRNIGALLPRLHQLCCMRPTRLRLFVAELCSWMLMTFTLYTDFVGEGL 300

Qy 301 YQGVPRAPGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
Db 301 YQGVPRAPGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360

Qy 361 APVVAAGATCLSHSVAVVTASAAITGFTFSAQLPILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APVVAAGATCLSHSVAVVTASAAITGFTFSAQLPILPYTLASLYHREKQVFLPKYRGDTGG 420

Qy 421 ASSEDSLMTSFLPGPKGAPFNGHVAGGSGLLPPPPALCGASACDVSVRVVVGTEA 480
Db 421 ASSEDSLMTSFLPGPKGAPFNGHVAGGSGLLPPPPALCGASACDVSVRVVVGTEA 480

Qy 481 RVVPRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYFATQ 540
Db 481 RVVPRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYFATQ 540

Qy 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

RESULT 4
US-09-352-616A-113
; Sequence 113, Application US/09352616A
; Patent No. 6395278
; GENERAL INFORMATION:
; APPLICANT: Dillon, Davin C.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang, Yugu
; APPLICANT: Xu, Jiangchun
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.427C8
; CURRENT APPLICATION NUMBER: US/09/352.616A
; CURRENT FILING DATE: 1999-07-13
; NUMBER OF SEQ ID NOS: 472
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553

Query Match      100.0%; Score 2861; DB 4; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.1e-270;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVQRLWVSRLRHRKQAQLLLVNLTFGLVCLAAAGITVPPLLELVGVVEEKFMVTLGIG 60
Db 1 MVQRLWVSRLRHRKQAQLLLVNLTFGLVCLAAAGITVPPLLELVGVVEEKFMVTLGIG 60

Qy 61 PVGLVCVPLIGSADHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
Db 61 PVGLVCVPLIGSADHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120

Qy 121 ELALLILGVGLDFCGQVCFPLEALLSDLPDPDCHQAYSVYAFMISLGGCLGYLLPA 180
Db 121 ELALLILGVGLDFCGQVCFPLEALLSDLPDPDCHQAYSVYAFMISLGGCLGYLLPA 180

Qy 181 IDWTSALAPYLGTQEECLFGLLTLLIFLTCAATLLVAEEAALGPTPEAGLSAPSLSPH 240
Db 181 IDWTSALAPYLGTQEECLFGLLTLLIFLTCAATLLVAEEAALGPTPEAGLSAPSLSPH 240

Qy 241 CCPCRLAFRNIGALLPRLHQLCCMRPTRLRLFVAELCSWMLMTFTLYTDFVGEGL 300
Db 241 CCPCRLAFRNIGALLPRLHQLCCMRPTRLRLFVAELCSWMLMTFTLYTDFVGEGL 300

Qy 301 YQGVPRAPGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
Db 301 YQGVPRAPGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360

Qy 361 APVVAAGATCLSHSVAVVTASAAITGFTFSAQLPILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APVVAAGATCLSHSVAVVTASAAITGFTFSAQLPILPYTLASLYHREKQVFLPKYRGDTGG 420

Qy 421 ASSEDSLMTSFLPGPKGAPFNGHVAGGSGLLPPPPALCGASACDVSVRVVVGTEA 480
Db 421 ASSEDSLMTSFLPGPKGAPFNGHVAGGSGLLPPPPALCGASACDVSVRVVVGTEA 480

Qy 481 RVVPRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYFATQ 540
Db 481 RVVPRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYFATQ 540

Qy 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

RESULT 5
US-09-602-877A-101
; Sequence 101, Application US/09602877A
; Patent No. 6432707
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.446C5
; CURRENT APPLICATION NUMBER: US/09/602.877A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 101
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-602-877A-101

Query Match      100.0%; Score 2861; DB 4; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.1e-270;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```


QY 1 MVQLWVSRLLRHRKAQALLNLTFTGLEVCIAAGITYVPPILLLEVGVEEKFTMTVLGIG 60
DB 1 MVQLWVSRLLRHRKAQALLNLTFTGLEVCIAAGITYVPPILLLEVGVEEKFTMTVLGIG 60
QY 61 PVGLVCPVLLGSASDHWGRGYGRRRPFPIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
DB 61 PVGLVCPVLLGSASDHWGRGYGRRRPFPIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
QY 121 ELALLILGVGLLDFCGOVCFPTLEALLSDLPDCHCRQAYSVYAFMISLGGCLGYLLPA 180
DB 121 ELALLILGVGLLDFCGOVCFPTLEALLSDLPDCHCRQAYSVYAFMISLGGCLGYLLPA 180
QY 181 IDWTSALAPYLGTQOECLFGLLTLFLTCAATLLVAEEAALGPTPEAGLSAPSLSPH 240
DB 181 IDWTSALAPYLGTQOECLFGLLTLFLTCAATLLVAEEAALGPTPEAGLSAPSLSPH 240
QY 241 CCPCRLAFRNIGALLPRHLQCCRMPTLRLRFLVAELCSMMALMTFTLYTDFVGEGL 300
DB 241 CCPCRLAFRNIGALLPRHLQCCRMPTLRLRFLVAELCSMMALMTFTLYTDFVGEGL 300
QY 301 YQGVPRAEPTGTEARRHYDEGRVMSGLGLFLOCAISLVFSLVMDRLVORFGTRAVYLASVA 360
DB 301 YQGVPRAEPTGTEARRHYDEGRVMSGLGLFLOCAISLVFSLVMDRLVORFGTRAVYLASVA 360
QY 361 APVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
DB 361 APVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGPKPGAPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGTEA 480
DB 421 ASSEDSLMTSFLPGPKPGAPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGTEA 480
QY 481 RVVPRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYFATQ 540
DB 481 RVVPRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYFATQ 540
QY 541 VVFDKSLAKYSA 553
DB 541 VVFDKSLAKYSA 553

RESULT 6

US-09-232-149A-113
; Sequence 113, Application US/09232149A
; Patent No. 6465611
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, David C.
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE
; FILE REFERENCE: 210121.427C6
; CURRENT APPLICATION NUMBER: US/09/232.149A
; CURRENT FILING DATE: 1999-01-15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-232-149A-113

Query Match 100.0%; Score 2861; DB 4; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.1e-270;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQLWVSRLLRHRKAQALLNLTFTGLEVCIAAGITYVPPILLLEVGVEEKFTMTVLGIG 60
DB 1 MVQLWVSRLLRHRKAQALLNLTFTGLEVCIAAGITYVPPILLLEVGVEEKFTMTVLGIG 60
QY 61 PVGLVCPVLLGSASDHWGRGYGRRRPFPIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120

DB 61 PVGLVCPVLLGSASDHWGRGYGRRRPFPIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
QY 121 ELALLILGVGLLDFCGOVCFPTLEALLSDLPDCHCRQAYSVYAFMISLGGCLGYLLPA 180
DB 121 ELALLILGVGLLDFCGOVCFPTLEALLSDLPDCHCRQAYSVYAFMISLGGCLGYLLPA 180
QY 181 IDWTSALAPYLGTQOECLFGLLTLFLTCAATLLVAEEAALGPTPEAGLSAPSLSPH 240
DB 181 IDWTSALAPYLGTQOECLFGLLTLFLTCAATLLVAEEAALGPTPEAGLSAPSLSPH 240
QY 241 CCPCRLAFRNIGALLPRHLQCCRMPTLRLRFLVAELCSMMALMTFTLYTDFVGEGL 300
DB 241 CCPCRLAFRNIGALLPRHLQCCRMPTLRLRFLVAELCSMMALMTFTLYTDFVGEGL 300
QY 301 YQGVPRAEPTGTEARRHYDEGRVMSGLGLFLOCAISLVFSLVMDRLVORFGTRAVYLASVA 360
DB 301 YQGVPRAEPTGTEARRHYDEGRVMSGLGLFLOCAISLVFSLVMDRLVORFGTRAVYLASVA 360
QY 361 APVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
DB 361 APVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGPKPGAPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGTEA 480
DB 421 ASSEDSLMTSFLPGPKPGAPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGTEA 480
QY 481 RVVPRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYFATQ 540
DB 481 RVVPRGICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYFATQ 540
QY 541 VVFDKSLAKYSA 553
DB 541 VVFDKSLAKYSA 553

RESULT 7

US-09-071-710-36
; Sequence 36, Application US/09071710
; Patent No. 6130043
; GENERAL INFORMATION:
; APPLICANT: BILLING-MEDEL, PATRICIA
; APPLICANT: COHEN, MAURICE
; APPLICANT: COLPITTS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
; APPLICANT: GORDON, JULIAN
; APPLICANT: GRANADOS, EDWARD N.
; APPLICANT: HODGES, STEVEN C.
; APPLICANT: KLASS, MICHAEL R.
; APPLICANT: KRATOCHVIL, JON D.
; APPLICANT: ROBERTS-RAPP, LISA
; APPLICANT: RUSSELL, JOHN C.
; APPLICANT: STROUPE, STEPHEN D.
; TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
; FOR DETECTING DISEASES OF THE PROSTATE
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,710
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/850,713

FILING DATE: 02-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6083.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 255 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6130043e
US-09-071-710-36

Query Match 45.0%; Score 1287; DB 3; Length 255;
Best Local Similarity 100.0%; Pred. No. 1.6e-117;
Matches 255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 299 GLYGVPRAEPTGTEARRHYDEGVGMGSLGLFLQCAISLVFSLVMDRLVQRFGRVAVLAS 358
DB 1 GLYGVPRAEPTGTEARRHYDEGVGMGSLGLFLQCAISLVFSLVMDRLVQRFGRVAVLAS 60

QY 359 VAAFPVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLYHREKOVFLPKYRGDT 418
DB 61 VAAFPVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLYHREKOVFLPKYRGDT 120

QY 419 GGASSEDLSMTSFLPGPKGAPFNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEP 478
DB 121 GGASSEDLSMTSFLPGPKGAPFNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEP 180

QY 479 EARVPGRGICLDLAIDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYFA 538
DB 181 EARVPGRGICLDLAIDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYFA 240

QY 539 TQVVFDSKDLAKYSA 553
DB 241 TQVVFDSKDLAKYSA 255

RESULT 8
US-09-525-397-36
Sequence 36, Application US/09525397
Patent No. 6252047
GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLASS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/525,397
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/071,710
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6083.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 255 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6252047e
US-09-525-397-36

Query Match 45.0%; Score 1287; DB 3; Length 255;
Best Local Similarity 100.0%; Pred. No. 1.6e-117;
Matches 255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 299 GLYGVPRAEPTGTEARRHYDEGVGMGSLGLFLQCAISLVFSLVMDRLVQRFGRVAVLAS 358
DB 1 GLYGVPRAEPTGTEARRHYDEGVGMGSLGLFLQCAISLVFSLVMDRLVQRFGRVAVLAS 60

QY 359 VAAFPVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLYHREKOVFLPKYRGDT 418
DB 61 VAAFPVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLYHREKOVFLPKYRGDT 120

QY 419 GGASSEDLSMTSFLPGPKGAPFNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEP 478
DB 121 GGASSEDLSMTSFLPGPKGAPFNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEP 180

QY 479 EARVPGRGICLDLAIDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYFA 538
DB 181 EARVPGRGICLDLAIDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIYFA 240

QY 539 TQVVFDSKDLAKYSA 553
DB 241 TQVVFDSKDLAKYSA 255

RESULT 9
US-09-439-313-571
Sequence 571, Application US/09439313
Patent No. 6329505
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan Louise
APPLICANT: Jiang Yuqi
APPLICANT: Reed, Steven G.
APPLICANT: Kalos, Michael
APPLICANT: Fanger, Gary
APPLICANT: Retter, Mark
APPLICANT: Solk, John
APPLICANT: Day, Craig
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.427C9
CURRENT APPLICATION NUMBER: US/09/439,313
CURRENT FILING DATE: 1999-11-12
NUMBER OF SEQ ID NOS: 575

RESULT 13

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2004, 17:03:52 ; Search time 39 Seconds
(without alignments)
2931.420 Million cell updates/sec

Title: US-09-593-793a-113

Perfect score: 2861

Sequence: 1 MVQRLWVSRLRHRKAQLL.....AIYFATQVDFKSLAKYSA 553

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 777136 seqs, 20673638 residues

Total number of hits satisfying chosen parameters: 777136

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA:*
- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT NEW PUB.pep:*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW PUB.pep:*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW PUB.pep:*
 - 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
 - 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW PUB.pep:*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
 - 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
 - 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW PUB.pep:*
 - 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
 - 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
 - 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
 - 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW PUB.pep:*
 - 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW PUB.pep:*
 - 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2861	100.0	553	9	US-09-745-288-101
2	2861	100.0	553	9	US-09-838-785-2
3	2861	100.0	553	9	US-09-759-143-113
4	2861	100.0	553	9	US-09-780-669-113
5	2861	100.0	553	9	US-09-030-606-113
6	2861	100.0	553	9	US-09-822-827-113
7	2861	100.0	553	9	US-09-115-453-113
8	2861	100.0	553	10	US-09-232-880-113
9	2861	100.0	553	10	US-09-895-793-113
10	2861	100.0	553	10	US-09-895-814-113
11	2861	100.0	553	12	US-10-144-678A-113
12	2861	100.0	553	12	US-10-005-907-13
13	2861	100.0	553	12	US-10-294-025-113
14	2861	100.0	553	12	US-10-295-027-548
15	2861	100.0	553	12	US-10-295-027-902

16	2861	100.0	553	14	US-10-012-896-113	Sequence 113, App
17	2861	100.0	553	15	US-10-010-940-113	Sequence 113, App
18	2861	100.0	1079	9	US-09-822-827-947	Sequence 947, App
19	2861	100.0	1079	10	US-09-895-793-947	Sequence 947, App
20	2601	90.9	710	15	US-10-296-770-4	Sequence 4, Appli
21	1696	59.3	359	9	US-09-822-827-974	Sequence 974, App
22	1696	59.3	359	10	US-09-895-793-974	Sequence 974, App
23	1677.5	58.6	530	15	US-10-296-770-5	Sequence 5, Appli
24	1517	53.0	305	12	US-10-144-678A-1029	Sequence 1029, Ap
25	1517	53.0	305	12	US-10-294-025-1029	Sequence 1029, Ap
26	1417.5	49.5	371	9	US-09-759-143-708	Sequence 708, App
27	1417.5	49.5	371	9	US-09-780-669-708	Sequence 708, App
28	1417.5	49.5	371	9	US-09-822-827-708	Sequence 708, App
29	1417.5	49.5	371	10	US-09-895-793-708	Sequence 708, App
30	1417.5	49.5	371	10	US-09-895-814-708	Sequence 708, App
31	1417.5	49.5	371	12	US-10-144-678A-708	Sequence 708, App
32	1417.5	49.5	371	12	US-10-294-025-708	Sequence 708, App
33	1417.5	49.5	371	14	US-10-012-896-708	Sequence 708, App
34	1416	49.5	371	12	US-09-833-245-852	Sequence 852, App
35	1403.5	49.1	400	9	US-09-759-143-852	Sequence 852, App
36	1403.5	49.1	400	9	US-09-780-669-852	Sequence 852, App
37	1403.5	49.1	400	9	US-09-822-827-852	Sequence 852, App
38	1403.5	49.1	400	10	US-09-895-793-852	Sequence 852, App
39	1403.5	49.1	400	10	US-09-895-814-852	Sequence 852, App
40	1403.5	49.1	400	12	US-10-144-678A-852	Sequence 852, App
41	1403.5	49.1	400	12	US-10-294-025-852	Sequence 852, App
42	1403.5	49.1	400	14	US-10-012-896-852	Sequence 852, App
43	1197	41.8	246	12	US-10-144-678A-1028	Sequence 1028, Ap
44	1197	41.8	246	12	US-10-294-025-1028	Sequence 1028, Ap
45	1014.5	35.5	326	12	US-10-005-907-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-09-745-288-101
; Sequence 101, Application US/09745288
; Patent No. US20010018058A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.446D1
; CURRENT APPLICATION NUMBER: US/09745.288
; CURRENT FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 101
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-745-288-101

Query Match	100.0%;	Score 2861;	DB 9;	Length 553;
Best Local Similarity	100.0%;	Pred. No. 2.2e+245;		
Matches	553;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
Qy	1	MVQRLWVSRLRHRKAQLLVNLITFGLEVCLAAGITVPPLLEVGVEEKFTMTVLIG	60	
Db	1	MVQRLWVSRLRHRKAQLLVNLITFGLEVCLAAGITVPPLLEVGVEEKFTMTVLIG	60	
Qy	61	PVLGLVCPVLIGSADHWGRYGRRRPFIVALSIGILLSLFLIPRAGWLAGLLCPDRPL	120	
Db	61	PVLGLVCPVLIGSADHWGRYGRRRPFIVALSIGILLSLFLIPRAGWLAGLLCPDRPL	120	
Qy	121	ELALLIIGVGLIDFCGVCTTPEALLSDFRDPDHCROAYSVYAFMISGGCGLYLLPA	180	
Db	121	ELALLIIGVGLIDFCGVCTTPEALLSDFRDPDHCROAYSVYAFMISGGCGLYLLPA	180	
Qy	181	IDWTSALAPYLGTOECLFGLLTIFLTCAATLLVAEEAALGPTPEAGLSAPLSLSPH	240	

```
Db 181 IDWTSALAPYLGTQBECLFGLLTLLIFLTCAATLLVAEEAALGPTPEAGLSAPLSPH 240
Qy 241 CCPCRLAFRNIGALLPRLHQLCCRMPTLRLRFVAELCSWALMTFTFLFYTDVGEGL 300
Db 241 CCPCRLAFRNIGALLPRLHQLCCRMPTLRLRFVAELCSWALMTFTFLFYTDVGEGL 300
Qy 301 YQGVPRAPGTEARRHYDEGRVMSGLGLFQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
Db 301 YQGVPRAPGTEARRHYDEGRVMSGLGLFQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
Qy 361 AFPVAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 AFPVAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Qy 421 ASSEDSLMTSFLPGPKGAPPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGTEA 480
Db 421 ASSEDSLMTSFLPGPKGAPPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGTEA 480
Qy 481 RVVPRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAIFYATQ 540
Db 481 RVVPRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAIFYATQ 540
Qy 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553
```

RESULT 2

```
US-09-838-785-2
; Sequence 2, Application US/09838785
; Patent No. US20020009455A1
; GENERAL INFORMATION:
; APPLICANT: Lin, Rick
; APPLICANT: Parkes, Debbie
; APPLICANT: Parry, Gordon
; APPLICANT: Schneider, Douglas
; APPLICANT: Steinbrecher, Renate
; APPLICANT: Van Heuit, Pam T
; APPLICANT: Wu, John
; TITLE OF INVENTION: DNA Encoding a No. US20020009455A1el PROST 03
; FILE REFERENCE: 51831AUSM1
; CURRENT APPLICATION NUMBER: US/09/838,785
; CURRENT FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/200,065
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-838-785-2
```

```
Query Match 100.0%; Score 2861; DB 9; Length 553;
Best Local Similarity 100.0%; Pred. No. 2.2e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVQRLWVSLRLRRKQAQLLLVNLITFGLEVCLAAGITVVPPLLEVGVEEKFTMWLGIG 60
Db 1 MVQRLWVSLRLRRKQAQLLLVNLITFGLEVCLAAGITVVPPLLEVGVEEKFTMWLGIG 60
Qy 61 PVGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
Db 61 PVGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
Qy 121 ELALLIIGVLLDPCGVCFPTLEALLSDIFRDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
Db 121 ELALLIIGVLLDPCGVCFPTLEALLSDIFRDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
Qy 181 IDWTSALAPYLGTQBECLFGLLTLLIFLTCAATLLVAEEAALGPTPEAGLSAPLSPH 240
```

```
Db 181 IDWTSALAPYLGTQBECLFGLLTLLIFLTCAATLLVAEEAALGPTPEAGLSAPLSPH 240
Qy 241 CCPCRLAFRNIGALLPRLHQLCCRMPTLRLRFVAELCSWALMTFTFLFYTDVGEGL 300
Db 241 CCPCRLAFRNIGALLPRLHQLCCRMPTLRLRFVAELCSWALMTFTFLFYTDVGEGL 300
Qy 301 YQGVPRAPGTEARRHYDEGRVMSGLGLFQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
Db 301 YQGVPRAPGTEARRHYDEGRVMSGLGLFQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
Qy 361 AFPVAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 AFPVAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Qy 421 ASSEDSLMTSFLPGPKGAPPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGTEA 480
Db 421 ASSEDSLMTSFLPGPKGAPPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGTEA 480
Qy 481 RVVPRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAIFYATQ 540
Db 481 RVVPRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAIFYATQ 540
Qy 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553
```

RESULT 3

```
US-09-759-143-113
; Sequence 113, Application US/09759143
; Patent No. US2002002248A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C23
; CURRENT APPLICATION NUMBER: US/09/759,143
; CURRENT FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-759-143-113
```

```
Query Match 100.0%; Score 2861; DB 9; Length 553;
Best Local Similarity 100.0%; Pred. No. 2.2e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVQRLWVSLRLRRKQAQLLLVNLITFGLEVCLAAGITVVPPLLEVGVEEKFTMWLGIG 60
Db 1 MVQRLWVSLRLRRKQAQLLLVNLITFGLEVCLAAGITVVPPLLEVGVEEKFTMWLGIG 60
Qy 61 PVGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
Db 61 PVGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDPRPL 120
```


QY 121 ELALLILGVLLDFCGQVCFPTLEALLSDLFDPDHCRCQAYSVYAFMISLGCGCLGYLLPA 180
DB 121 ELALLILGVLLDFCGQVCFPTLEALLSDLFDPDHCRCQAYSVYAFMISLGCGCLGYLLPA 180
QY 181 IDWDTSAAPYLGTOEECLFGLLTLLIPLTCTVAATLLVAEEAALGTEPAEGLSAPLSPH 240
DB 181 IDWDTSAAPYLGTOEECLFGLLTLLIPLTCTVAATLLVAEEAALGTEPAEGLSAPLSPH 240
QY 241 CCPCRLAFRNILGALLPRLHQLCCRMPTLRLRFVAELCSWMLMTFTLFTYDFVGEGL 300
DB 241 CCPCRLAFRNILGALLPRLHQLCCRMPTLRLRFVAELCSWMLMTFTLFTYDFVGEGL 300
QY 301 YQGVPRAPGTEARRHYDEGVMSGLGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
DB 301 YQGVPRAPGTEARRHYDEGVMSGLGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
QY 361 APPVAAGATCLSHSVAVVTASAAITGFTFSAQILPYTLASLYHREKQVFLPKYRGDTGG 420
DB 361 APPVAAGATCLSHSVAVVTASAAITGFTFSAQILPYTLASLYHREKQVFLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGKPGAPFNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
DB 421 ASSEDSLMTSFLPGKPGAPFNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
QY 481 RVVPRGICDLAILDLSAFLLSQVAPSLFMSI1VOLQSQTAYMVSAAGLGLVAIYFATQ 540
DB 481 RVVPRGICDLAILDLSAFLLSQVAPSLFMSI1VOLQSQTAYMVSAAGLGLVAIYFATQ 540
QY 541 VVFDKSLAKYSA 553
DB 541 VVFDKSLAKYSA 553

RESULT 4

US-09-780-669-113
; Sequence 113, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Jennifer L.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yuqi
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Retter, John W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C24
CURRENT APPLICATION NUMBER: US/09/780.669
CURRENT FILING DATE: 2001-02-09
NUMBER OF SEQ ID NOS: 943
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 113
LENGTH: 553
TYPE: PRT
ORGANISM: Homo sapien

US-09-780-669-113

Query Match

100.0%; Score 2861; DB 9; Length 553;

Best Local Similarity 100.0%; Pred. No. 2.2e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVQGLWYSRLRHRKAQOLLVNLITFGLVLEVCIAAGITVVPPLLELVGVEEKFTMWLGIG 60
DB 1 MVQGLWYSRLRHRKAQOLLVNLITFGLVLEVCIAAGITVVPPLLELVGVEEKFTMWLGIG 60
QY 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLI PRAGWLAGLCPDPRPL 120
DB 61 PVLGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLI PRAGWLAGLCPDPRPL 120
QY 121 ELALLILGVLLDFCGQVCFPTLEALLSDLFDPDHCRCQAYSVYAFMISLGCGCLGYLLPA 180
DB 121 ELALLILGVLLDFCGQVCFPTLEALLSDLFDPDHCRCQAYSVYAFMISLGCGCLGYLLPA 180
QY 181 IDWDTSAAPYLGTOEECLFGLLTLLIPLTCTVAATLLVAEEAALGTEPAEGLSAPLSPH 240
DB 181 IDWDTSAAPYLGTOEECLFGLLTLLIPLTCTVAATLLVAEEAALGTEPAEGLSAPLSPH 240
QY 241 CCPCRLAFRNILGALLPRLHQLCCRMPTLRLRFVAELCSWMLMTFTLFTYDFVGEGL 300
DB 241 CCPCRLAFRNILGALLPRLHQLCCRMPTLRLRFVAELCSWMLMTFTLFTYDFVGEGL 300
QY 301 YQGVPRAPGTEARRHYDEGVMSGLGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
DB 301 YQGVPRAPGTEARRHYDEGVMSGLGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
QY 361 APPVAAGATCLSHSVAVVTASAAITGFTFSAQILPYTLASLYHREKQVFLPKYRGDTGG 420
DB 361 APPVAAGATCLSHSVAVVTASAAITGFTFSAQILPYTLASLYHREKQVFLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGKPGAPFNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
DB 421 ASSEDSLMTSFLPGKPGAPFNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
QY 481 RVVPRGICDLAILDLSAFLLSQVAPSLFMSI1VOLQSQTAYMVSAAGLGLVAIYFATQ 540
DB 481 RVVPRGICDLAILDLSAFLLSQVAPSLFMSI1VOLQSQTAYMVSAAGLGLVAIYFATQ 540
QY 541 VVFDKSLAKYSA 553
DB 541 VVFDKSLAKYSA 553

RESULT 5

US-09-030-606-113
; Sequence 113, Application US/09030606
; Patent No. US20020081580A1
; GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
TITLE OF INVENTION: COMPOUNDS FOR IMMUNODIAGNOSIS OF PROSTATE CANCER AND METHODS F
NUMBER OF SEQUENCES: 224
CORRESPONDENCE ADDRESS:
ADDRESS: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/030,606
FILING DATE: 25-FEB-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.428C3

TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 113:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 553 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; US-09-030-606-113

Query Match 100.0%; Score 2861; DB 9; Length 553;
Best Local Similarity 100.0%; Pred. No. 2.2e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLRRHKAQLLLVNLITFGLVCLAAAGITYVPPLLLLEVGVBEKFTMWLGIG 60
Db 1 MVQRLWVSRLRRHKAQLLLVNLITFGLVCLAAAGITYVPPLLLLEVGVBEKFTMWLGIG 60

QY 61 PVGLVCPVLLGSADHWGRGRRRPFIFWALSGLISLFLIPRAGWLAGLCPDPRPL 120
Db 61 PVGLVCPVLLGSADHWGRGRRRPFIFWALSGLISLFLIPRAGWLAGLCPDPRPL 120

QY 121 ELALLILGVGLLDFCGQVCFPLEALLSDLFDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
Db 121 ELALLILGVGLLDFCGQVCFPLEALLSDLFDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180

QY 181 IDWTSALAPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
Db 181 IDWTSALAPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240

QY 241 CCPCRLAFRNIGALLPRLHQLCCRPRTLRLFLVAELCSNMALMTFTFLFYDFVGEGL 300
Db 241 CCPCRLAFRNIGALLPRLHQLCCRPRTLRLFLVAELCSNMALMTFTFLFYDFVGEGL 300

QY 301 YQGVPRAEPTGTEARRHYDEGRVMSLGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
Db 301 YQGVPRAEPTGTEARRHYDEGRVMSLGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360

QY 361 APFVAAGATCLSHSVAVVTASAAITGTFPSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APFVAAGATCLSHSVAVVTASAAITGTFPSALQILPYTLASLYHREKQVFLPKYRGDTGG 420

QY 421 ASSDSLMTSFLPGPKGAPPFNGHVAGSGGLPPPPALCGASACDVSVRVVVGEPTEA 480
Db 421 ASSDSLMTSFLPGPKGAPPFNGHVAGSGGLPPPPALCGASACDVSVRVVVGEPTEA 480

QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAAGLGLVAIYPATQ 540
Db 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAAGLGLVAIYPATQ 540

QY 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

RESULT 6
US-09-822-827-113
; Sequence 113, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822,827
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113

Query Match 100.0%; Score 2861; DB 9; Length 553;
Best Local Similarity 100.0%; Pred. No. 2.2e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLRRHKAQLLLVNLITFGLVCLAAAGITYVPPLLLLEVGVBEKFTMWLGIG 60
Db 1 MVQRLWVSRLRRHKAQLLLVNLITFGLVCLAAAGITYVPPLLLLEVGVBEKFTMWLGIG 60

QY 61 PVGLVCPVLLGSADHWGRGRRRPFIFWALSGLISLFLIPRAGWLAGLCPDPRPL 120
Db 61 PVGLVCPVLLGSADHWGRGRRRPFIFWALSGLISLFLIPRAGWLAGLCPDPRPL 120

QY 121 ELALLILGVGLLDFCGQVCFPLEALLSDLFDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
Db 121 ELALLILGVGLLDFCGQVCFPLEALLSDLFDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180

QY 181 IDWTSALAPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
Db 181 IDWTSALAPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240

QY 241 CCPCRLAFRNIGALLPRLHQLCCRPRTLRLFLVAELCSNMALMTFTFLFYDFVGEGL 300
Db 241 CCPCRLAFRNIGALLPRLHQLCCRPRTLRLFLVAELCSNMALMTFTFLFYDFVGEGL 300

QY 301 YQGVPRAEPTGTEARRHYDEGRVMSLGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
Db 301 YQGVPRAEPTGTEARRHYDEGRVMSLGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360

QY 361 APFVAAGATCLSHSVAVVTASAAITGTFPSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APFVAAGATCLSHSVAVVTASAAITGTFPSALQILPYTLASLYHREKQVFLPKYRGDTGG 420

QY 421 ASSDSLMTSFLPGPKGAPPFNGHVAGSGGLPPPPALCGASACDVSVRVVVGEPTEA 480
Db 421 ASSDSLMTSFLPGPKGAPPFNGHVAGSGGLPPPPALCGASACDVSVRVVVGEPTEA 480

QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAAGLGLVAIYPATQ 540
Db 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAAGLGLVAIYPATQ 540

QY 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

RESULT 7
US-09-115-453-113
; Sequence 113, Application US/09115453B
; Patent No. US20020090372A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND
; FILE REFERENCE: 210121.427C4
; CURRENT APPLICATION NUMBER: US/09/115,453B
; CURRENT FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113

Query Match 100.0%; Score 2861; DB 9; Length 553;
Best Local Similarity 100.0%; Pred. No. 2.2e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLRRHKAQLLLVNLITFGLVCLAAAGITYVPPLLLLEVGVBEKFTMWLGIG 60
Db 1 MVQRLWVSRLRRHKAQLLLVNLITFGLVCLAAAGITYVPPLLLLEVGVBEKFTMWLGIG 60

QY 61 PVGLVCPVLLGSADHWGRGRRRPFIFWALSGLISLFLIPRAGWLAGLCPDPRPL 120
Db 61 PVGLVCPVLLGSADHWGRGRRRPFIFWALSGLISLFLIPRAGWLAGLCPDPRPL 120

QY 121 ELALLILGVGLLDFCGQVCFPLEALLSDLFDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
Db 121 ELALLILGVGLLDFCGQVCFPLEALLSDLFDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180

QY 181 IDWTSALAPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
Db 181 IDWTSALAPYLGTQBECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240

QY 241 CCPCRLAFRNIGALLPRLHQLCCRPRTLRLFLVAELCSNMALMTFTFLFYDFVGEGL 300
Db 241 CCPCRLAFRNIGALLPRLHQLCCRPRTLRLFLVAELCSNMALMTFTFLFYDFVGEGL 300

QY 301 YQGVPRAEPTGTEARRHYDEGRVMSLGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360
Db 301 YQGVPRAEPTGTEARRHYDEGRVMSLGLFLQCAISLVFSLVMDRLVQRFCTRAVYLASVA 360

QY 361 APFVAAGATCLSHSVAVVTASAAITGTFPSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APFVAAGATCLSHSVAVVTASAAITGTFPSALQILPYTLASLYHREKQVFLPKYRGDTGG 420

QY 421 ASSDSLMTSFLPGPKGAPPFNGHVAGSGGLPPPPALCGASACDVSVRVVVGEPTEA 480
Db 421 ASSDSLMTSFLPGPKGAPPFNGHVAGSGGLPPPPALCGASACDVSVRVVVGEPTEA 480

QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAAGLGLVAIYPATQ 540
Db 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAAGLGLVAIYPATQ 540

QY 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

```
Qy 1 MVQRLWSRLRHRKAQALLVNLITFGLVCLAAAGITVPPLLEVGVEEKFTMTVLGIG 60
Db 1 MVQRLWSRLRHRKAQALLVNLITFGLVCLAAAGITVPPLLEVGVEEKFTMTVLGIG 60
Qy 61 PVGLVCPVLGSGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLCPDRPL 120
Db 61 PVGLVCPVLGSGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLCPDRPL 120
Qy 121 ELALLILGVLLDFCGQVCTPPEALLSDLPDHPDHCQAYSVYAFMISLGGCLGYLLPA 180
Db 121 ELALLILGVLLDFCGQVCTPPEALLSDLPDHPDHCQAYSVYAFMISLGGCLGYLLPA 180
Qy 181 IDWTSALAPYLGTOEBCFLGLLTLIFLTCVAATLLVAEEAALGPTPEAGLSAPSLSPH 240
Db 181 IDWTSALAPYLGTOEBCFLGLLTLIFLTCVAATLLVAEEAALGPTPEAGLSAPSLSPH 240
Qy 241 CCPCRLAFRNIGALLPRLHQLCCRPRTLRLLFVAELCSWMLMTFTLYTDFVGEGL 300
Db 241 CCPCRLAFRNIGALLPRLHQLCCRPRTLRLLFVAELCSWMLMTFTLYTDFVGEGL 300
Qy 301 YQGVPRAPGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
Db 301 YQGVPRAPGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
Qy 361 APVVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APVVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Qy 421 ASSEDSLMTSFLPGPKGAPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGTEA 480
Db 421 ASSEDSLMTSFLPGPKGAPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGTEA 480
Qy 481 RVVPGRGICLDLAILDSAFLLSOVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAIYATQ 540
Db 481 RVVPGRGICLDLAILDSAFLLSOVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAIYATQ 540
Qy 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

RESULT 8
US-09-232-880-113
; Sequence 113, Application US/09232880
; Publication No. US20020182596A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNODIAGNOSIS OF
; TITLE OF INVENTION: PROSTATE CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.428C6
; CURRENT APPLICATION NUMBER: US/09/232.880
; CURRENT FILING DATE: 1999-01-15
; NUMBER OF SEQ ID NOS: 338
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-232-880-113

Query Match 100.0%; Score 2861; DB 10; Length 553;
Best Local Similarity 100.0%; Pred. No. 2.2e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVQRLWSRLRHRKAQALLVNLITFGLVCLAAAGITVPPLLEVGVEEKFTMTVLGIG 60
Db 1 MVQRLWSRLRHRKAQALLVNLITFGLVCLAAAGITVPPLLEVGVEEKFTMTVLGIG 60
Qy 61 PVGLVCPVLGSGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLCPDRPL 120
Db 61 PVGLVCPVLGSGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLCPDRPL 120
```

```
Db 61 PVGLVCPVLGSGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLCPDRPL 120
Qy 121 ELALLILGVLLDFCGQVCTPPEALLSDLPDHPDHCQAYSVYAFMISLGGCLGYLLPA 180
Db 121 ELALLILGVLLDFCGQVCTPPEALLSDLPDHPDHCQAYSVYAFMISLGGCLGYLLPA 180
Qy 181 IDWTSALAPYLGTOEBCFLGLLTLIFLTCVAATLLVAEEAALGPTPEAGLSAPSLSPH 240
Db 181 IDWTSALAPYLGTOEBCFLGLLTLIFLTCVAATLLVAEEAALGPTPEAGLSAPSLSPH 240
Qy 241 CCPCRLAFRNIGALLPRLHQLCCRPRTLRLLFVAELCSWMLMTFTLYTDFVGEGL 300
Db 241 CCPCRLAFRNIGALLPRLHQLCCRPRTLRLLFVAELCSWMLMTFTLYTDFVGEGL 300
Qy 301 YQGVPRAPGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
Db 301 YQGVPRAPGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
Qy 361 APVVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APVVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Qy 421 ASSEDSLMTSFLPGPKGAPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGTEA 480
Db 421 ASSEDSLMTSFLPGPKGAPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGTEA 480
Qy 481 RVVPGRGICLDLAILDSAFLLSOVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAIYATQ 540
Db 481 RVVPGRGICLDLAILDSAFLLSOVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAIYATQ 540
Qy 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

RESULT 9
US-09-895-793-113
; Sequence 113, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedrick, Thomas S.
; APPLICANT: Carter, Darick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C2
; CURRENT APPLICATION NUMBER: US/09/895.793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
```

US-09-895-793-113

Query Match 100.0%; Score 2861; DB 10; Length 553;
Best Local Similarity 100.0%; Pred. No. 2.2e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVORLWVSRLLRRHKAQALLVNLITFTGLEVCCLAAGITVVPPLLEVGVEEKFMTWLVIG 60
Db 1 MVORLWVSRLLRRHKAQALLVNLITFTGLEVCCLAAGITVVPPLLEVGVEEKFMTWLVIG 60

Qy 61 PVGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDRPL 120
Db 61 PVGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDRPL 120

Qy 121 ELALLILGVGLDFCGVCFPTLEALLSDLFDPDHCRCQAYSVVAFMISGLGCLGYLLPA 180
Db 121 ELALLILGVGLDFCGVCFPTLEALLSDLFDPDHCRCQAYSVVAFMISGLGCLGYLLPA 180

Qy 181 IDWTSALAPYLGTQBECLFGLLTLLITCTVAATLLVAEEAALGPTPEAGLSAPLSPH 240
Db 181 IDWTSALAPYLGTQBECLFGLLTLLITCTVAATLLVAEEAALGPTPEAGLSAPLSPH 240

Qy 241 CCPCRLAFRNIGALLPRHLQCCRMPTLRLFLVAELCSWMAALMTFTLYTDFVGEGL 300
Db 241 CCPCRLAFRNIGALLPRHLQCCRMPTLRLFLVAELCSWMAALMTFTLYTDFVGEGL 300

Qy 301 YQGVPRAEPTGTEARRHYDEGRVMSGLGLFQCAISLVFSLVMDRLVQRFQTRAVYLA 360
Db 301 YQGVPRAEPTGTEARRHYDEGRVMSGLGLFQCAISLVFSLVMDRLVQRFQTRAVYLA 360

Qy 361 APVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420

Qy 421 ASSDSLTMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTA 480
Db 421 ASSDSLTMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTA 480

Qy 481 RVVPRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAIYPATQ 540
Db 481 RVVPRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAIYPATQ 540

Qy 541 VWFKSDLAKEYSA 553
Db 541 VWFKSDLAKEYSA 553

RESULT 10

US-09-895-814-113
; Sequence 113, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stoik, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darriack
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa

; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCES: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-895-814-113

Query Match 100.0%; Score 2861; DB 10; Length 553;
Best Local Similarity 100.0%; Pred. No. 2.2e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVORLWVSRLLRRHKAQALLVNLITFTGLEVCCLAAGITVVPPLLEVGVEEKFMTWLVIG 60
Db 1 MVORLWVSRLLRRHKAQALLVNLITFTGLEVCCLAAGITVVPPLLEVGVEEKFMTWLVIG 60

Qy 61 PVGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDRPL 120
Db 61 PVGLVCVPLLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLLCPDRPL 120

Qy 121 ELALLILGVGLDFCGVCFPTLEALLSDLFDPDHCRCQAYSVVAFMISGLGCLGYLLPA 180
Db 121 ELALLILGVGLDFCGVCFPTLEALLSDLFDPDHCRCQAYSVVAFMISGLGCLGYLLPA 180

Qy 181 IDWTSALAPYLGTQBECLFGLLTLLITCTVAATLLVAEEAALGPTPEAGLSAPLSPH 240
Db 181 IDWTSALAPYLGTQBECLFGLLTLLITCTVAATLLVAEEAALGPTPEAGLSAPLSPH 240

Qy 241 CCPCRLAFRNIGALLPRHLQCCRMPTLRLFLVAELCSWMAALMTFTLYTDFVGEGL 300
Db 241 CCPCRLAFRNIGALLPRHLQCCRMPTLRLFLVAELCSWMAALMTFTLYTDFVGEGL 300

Qy 301 YQGVPRAEPTGTEARRHYDEGRVMSGLGLFQCAISLVFSLVMDRLVQRFQTRAVYLA 360
Db 301 YQGVPRAEPTGTEARRHYDEGRVMSGLGLFQCAISLVFSLVMDRLVQRFQTRAVYLA 360

Qy 361 APVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420

Qy 421 ASSDSLTMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTA 480
Db 421 ASSDSLTMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTA 480

Qy 481 RVVPRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAIYPATQ 540
Db 481 RVVPRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAIYPATQ 540

Qy 541 VWFKSDLAKEYSA 553
Db 541 VWFKSDLAKEYSA 553

RESULT 11

US-10-144-678A-113
; Sequence 113, Application US/10144678A
; Publication No. US20030157089A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.

APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedwick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A. W.
APPLICANT: Hepler, William T.
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals y de Bassols, Carlota
APPLICANT: Foy, Teresa M.
APPLICANT: Watanabe, Yoshihiro
APPLICANT: Deng, Ta
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C28
CURRENT APPLICATION NUMBER: US/10/144,678A
CURRENT FILING DATE: 2002-08-12
NUMBER OF SEQ ID NOS: 1033
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 113
LENGTH: 553
TYPE: PRT
ORGANISM: Homo sapiens
US-10-144-678A-113

Query Match 100.0%; Score 2861; DB 12; Length 553;
Best Local Similarity 100.0%; Pred. No. 2.2e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MVQRLWVSRLRHRKAQLLVNLITFGLEVCLAAGITVVPPLLEVGVEEKFTMWLIG 60
Db 1 MVQRLWVSRLRHRKAQLLVNLITFGLEVCLAAGITVVPPLLEVGVEEKFTMWLIG 60
Qy 61 PVLGLVCPVLLGASDHWGRYGRRRPFIWALSGLILLSLFLIPRAGWLAGLLCPDRPL 120
Db 61 PVLGLVCPVLLGASDHWGRYGRRRPFIWALSGLILLSLFLIPRAGWLAGLLCPDRPL 120
Qy 121 ELALLILGVGLDFCGQVCFPLEALLSDLPDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
Db 121 ELALLILGVGLDFCGQVCFPLEALLSDLPDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
Qy 181 IDWTSALAPYLGTQEECLFGLTLITLTCVAATLLVAEEAALGPTPEAGLSAPSLSPH 240
Db 181 IDWTSALAPYLGTQEECLFGLTLITLTCVAATLLVAEEAALGPTPEAGLSAPSLSPH 240
Qy 241 CCPCRLAFRNLCALLPRLHQLCCMRPTRLRLFVAELCSWMLMTFTFLFYTDVFGGL 300
Db 241 CCPCRLAFRNLCALLPRLHQLCCMRPTRLRLFVAELCSWMLMTFTFLFYTDVFGGL 300
Qy 301 YQGVPRAPGTEARRHYDEGVRMGSGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
Db 301 YQGVPRAPGTEARRHYDEGVRMGSGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
Qy 361 APVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Qy 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTA 480
Db 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTA 480
Qy 481 RVVPRGICLDLAILDSAFLLSQVAPSLFMGSIVOLQSQSVTAYMVSAAGLGLVAIYPATQ 540
Db 481 RVVPRGICLDLAILDSAFLLSQVAPSLFMGSIVOLQSQSVTAYMVSAAGLGLVAIYPATQ 540
Qy 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

RESULT 13
US-10-294-025-113
; Sequence 113, Application US/10294025
; Publication No. US20030185830A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun

RESULT 12
US-10-005-907-13
; Sequence 13, Application US/10005907
; Publication No. US20030166881A1
; GENERAL INFORMATION:
; APPLICANT: Union Chimique Belge, S.A.
; APPLICANT: No. US20030166881A1ka, Karl
; APPLICANT: Pirozzi, Gregory
; APPLICANT: Einstein, Richard
; TITLE OF INVENTION: NOVEL GENES ASSOCIATED WITH ALLERGIC HYPERSENSITIVITY AND MAST CE
; TITLE OF INVENTION: ACTIVATION
; FILE REFERENCE: 053529-5005
; CURRENT APPLICATION NUMBER: US/10/005,907
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-005-907-13

Query Match 100.0%; Score 2861; DB 12; Length 553;
Best Local Similarity 100.0%; Pred. No. 2.2e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MVQRLWVSRLRHRKAQLLVNLITFGLEVCLAAGITVVPPLLEVGVEEKFTMWLIG 60
Db 1 MVQRLWVSRLRHRKAQLLVNLITFGLEVCLAAGITVVPPLLEVGVEEKFTMWLIG 60
Qy 61 PVLGLVCPVLLGASDHWGRYGRRRPFIWALSGLILLSLFLIPRAGWLAGLLCPDRPL 120
Db 61 PVLGLVCPVLLGASDHWGRYGRRRPFIWALSGLILLSLFLIPRAGWLAGLLCPDRPL 120
Qy 121 ELALLILGVGLDFCGQVCFPLEALLSDLPDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
Db 121 ELALLILGVGLDFCGQVCFPLEALLSDLPDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
Qy 181 IDWTSALAPYLGTQEECLFGLTLITLTCVAATLLVAEEAALGPTPEAGLSAPSLSPH 240
Db 181 IDWTSALAPYLGTQEECLFGLTLITLTCVAATLLVAEEAALGPTPEAGLSAPSLSPH 240
Qy 241 CCPCRLAFRNLCALLPRLHQLCCMRPTRLRLFVAELCSWMLMTFTFLFYTDVFGGL 300
Db 241 CCPCRLAFRNLCALLPRLHQLCCMRPTRLRLFVAELCSWMLMTFTFLFYTDVFGGL 300
Qy 301 YQGVPRAPGTEARRHYDEGVRMGSGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
Db 301 YQGVPRAPGTEARRHYDEGVRMGSGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
Qy 361 APVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 APVAAGATCLSHSVAVVTASAAITGFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Qy 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTA 480
Db 421 ASSEDSLMTSFLPGPKGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTA 480
Qy 481 RVVPRGICLDLAILDSAFLLSQVAPSLFMGSIVOLQSQSVTAYMVSAAGLGLVAIYPATQ 540
Db 481 RVVPRGICLDLAILDSAFLLSQVAPSLFMGSIVOLQSQSVTAYMVSAAGLGLVAIYPATQ 540
Qy 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

RESULT 13
US-10-294-025-113
; Sequence 113, Application US/10294025
; Publication No. US20030185830A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun

```

; APPLICANT: Stolk, John A.
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C29
; CURRENT APPLICATION NUMBER: US/10/294,025
; CURRENT FILING DATE: 2002-11-12
; NUMBER OF SEQ ID NOS: 1038
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 113
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-294-025-113

Query Match      100.0%; Score 2861; DB 12; Length 553;
Best Local Similarity 100.0%; Pred. No. 2.2e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVQRLWSRLRHRKAQLLLVNTLTFGLVCLAAAGITYVPPLLEVGVEEKFMTVLGIG 60
Db 1 MVQRLWSRLRHRKAQLLLVNTLTFGLVCLAAAGITYVPPLLEVGVEEKFMTVLGIG 60
Qy 61 PVLGLVCVPLLGASDHWGRYGRRRPFIFWALSIGILLSLFLIPRAGWLAGLLCPDRPL 120
Db 61 PVLGLVCVPLLGASDHWGRYGRRRPFIFWALSIGILLSLFLIPRAGWLAGLLCPDRPL 120
Qy 121 ELALLILGVLLDFCGQVCFPLEALLSDLFDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
Db 121 ELALLILGVLLDFCGQVCFPLEALLSDLFDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
Qy 181 IDWTSALAPYLGTQEECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
Db 181 IDWTSALAPYLGTQEECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
Qy 241 CCPCRARLAFRNLGALLPRHLQCCRMPTLRLLFVAELCSWMALMTFTFLFYDFVGEGL 300
Db 241 CCPCRARLAFRNLGALLPRHLQCCRMPTLRLLFVAELCSWMALMTFTFLFYDFVGEGL 300
Qy 301 YQGVPRAEPTGEARRHYDEGVNMGSLGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
Db 301 YQGVPRAEPTGEARRHYDEGVNMGSLGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
Qy 361 AFPVAAGATCLSHSVAVVTASAALTGTFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 AFPVAAGATCLSHSVAVVTASAALTGTFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Qy 421 ASSEDSLMTSFLPGPKGAPPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Db 421 ASSEDSLMTSFLPGPKGAPPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Qy 481 RVVPGRGICLDLALDLSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAIYPATQ 540
Db 481 RVVPGRGICLDLALDLSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAIYPATQ 540

RESULT 14
US-10-295-027-548
; Sequence 548, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glyne, Richard
; APPLICANT: Hezezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
```

```

; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 548
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-548
```

```

Query Match      100.0%; Score 2861; DB 12; Length 553;
Best Local Similarity 100.0%; Pred. No. 2.2e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVQRLWSRLRHRKAQLLLVNTLTFGLVCLAAAGITYVPPLLEVGVEEKFMTVLGIG 60
Db 1 MVQRLWSRLRHRKAQLLLVNTLTFGLVCLAAAGITYVPPLLEVGVEEKFMTVLGIG 60
Qy 61 PVLGLVCVPLLGASDHWGRYGRRRPFIFWALSIGILLSLFLIPRAGWLAGLLCPDRPL 120
Db 61 PVLGLVCVPLLGASDHWGRYGRRRPFIFWALSIGILLSLFLIPRAGWLAGLLCPDRPL 120
Qy 121 ELALLILGVLLDFCGQVCFPLEALLSDLFDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
Db 121 ELALLILGVLLDFCGQVCFPLEALLSDLFDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
Qy 181 IDWTSALAPYLGTQEECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
Db 181 IDWTSALAPYLGTQEECLFGLLTFLITCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
Qy 241 CCPCRARLAFRNLGALLPRHLQCCRMPTLRLLFVAELCSWMALMTFTFLFYDFVGEGL 300
Db 241 CCPCRARLAFRNLGALLPRHLQCCRMPTLRLLFVAELCSWMALMTFTFLFYDFVGEGL 300
Qy 301 YQGVPRAEPTGEARRHYDEGVNMGSLGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
Db 301 YQGVPRAEPTGEARRHYDEGVNMGSLGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
Qy 361 AFPVAAGATCLSHSVAVVTASAALTGTFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 AFPVAAGATCLSHSVAVVTASAALTGTFTFSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
Qy 421 ASSEDSLMTSFLPGPKGAPPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Db 421 ASSEDSLMTSFLPGPKGAPPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Qy 481 RVVPGRGICLDLALDLSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAIYPATQ 540
```

Db 481 RVVPGRGICLDLALDLSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIFYATQ 540
Qy 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

RESULT 15

US-10-295-027-902
; Sequence 902, Application US/10295027
; Publication No. US2003023350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 902
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-902

Query Match 100.0%; Score 2861; DB 12; Length 553;
Best Local Similarity 100.0%; Pred. No. 2.2e-245;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MVQRLWVSRLRHRKAQLLLVNLLTFGLVCLAAAGITVPPPLLEVGVEEKFTMTVLIG 60
Db 1 MVQRLWVSRLRHRKAQLLLVNLLTFGLVCLAAAGITVPPPLLEVGVEEKFTMTVLIG 60
Qy 61 PVLGLVCVPLIGSADHWGRYGRRRPFIWALSIGLLSLFLIPRAGWLAGLCPDRPL 120
Db 61 PVLGLVCVPLIGSADHWGRYGRRRPFIWALSIGLLSLFLIPRAGWLAGLCPDRPL 120
Qy 121 ELALLILGVLLDFCGVCFPTPLALLSDLPDRDPDHCROAYSVYAFMISLGGCLGYLLPA 180
Db 121 ELALLILGVLLDFCGVCFPTPLALLSDLPDRDPDHCROAYSVYAFMISLGGCLGYLLPA 180

Qy 181 IDWDTSLAPYLGTQBERCLFGLLTLLIFLTCAATLLVAEEAALGPTPEAGLSAPSLSPH 240
Db 181 IDWDTSLAPYLGTQBERCLFGLLTLLIFLTCAATLLVAEEAALGPTPEAGLSAPSLSPH 240
Qy 241 CCPCRARLAFNLCGALLPRLHOLCCRPRTLRLPVAELCSMMALMTFTLFTYDFVGEGL 300
Db 241 CCPCRARLAFNLCGALLPRLHOLCCRPRTLRLPVAELCSMMALMTFTLFTYDFVGEGL 300
Qy 301 YQGVPRAEPTGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
Db 301 YQGVPRAEPTGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
Qy 361 AFPVAAGATCLSHSVAVVTASAAITGFTFSAQLPYTLASLYHREKQVFLPKYRGDTGG 420
Db 361 AFPVAAGATCLSHSVAVVTASAAITGFTFSAQLPYTLASLYHREKQVFLPKYRGDTGG 420
Qy 421 ASSEDSLMTSFLPGPKPGAPPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Db 421 ASSEDSLMTSFLPGPKPGAPPPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
Qy 481 RVVPGRGICLDLALDLSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIFYATQ 540
Db 481 RVVPGRGICLDLALDLSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAIFYATQ 540
Qy 541 VVFDKSDLAKYSA 553
Db 541 VVFDKSDLAKYSA 553

Search completed: January 25, 2004, 17:07:19
Job time : 41 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2004, 17:01:42 ; Search time 21 Seconds
(without alignments)
2532.442 Million cell updates/sec

Title: US-09-593-793A-113

Perfect score: 2861

Sequence: 1 MVQRLWVSRLLRHRKAQLL.....AIYFATQVVFDDSLAKYSA 553

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_76.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	347.5	12.1	515	2 T14340	sucrose-proton tra
2	337.5	11.8	523	2 T12198	sucrose transport
3	330	11.5	516	2 JQ2369	sucrose transport
4	328.5	11.5	594	2 G84441	probable sucrose-p
5	324.5	11.3	525	2 S28052	sucrose transport
6	323.5	11.3	512	2 S38196	sucrose transport
7	323.5	11.3	512	2 G96360	probable sucrose-p
8	317.5	11.1	507	2 S48789	sucrose transport
9	311	10.9	533	2 S43142	sucrose transport
10	305.5	10.7	501	2 T14339	sucrose-proton tra
11	297.5	10.4	428	2 S48788	sucrose transport
12	294	10.3	510	2 S38657	sucrose transport
13	294	10.3	523	2 S21114	sucrose-proton sym
14	292.5	10.2	512	2 F96741	probable sucrose t
15	292	10.2	537	2 T02982	probable sucrose t
16	286	10.0	474	2 A86234	hypothetical prote
17	281.5	9.8	491	2 E96691	probable sucrose-p
18	279.5	9.8	513	2 S38197	sucrose transport
19	276.5	9.7	492	2 A84520	probable sucrose-p
20	273	9.5	503	2 S23777	sucrose transport
21	235.5	8.2	553	2 T38541	probable sucrose c
22	191	6.7	452	2 F75217	hypothetical prote
23	187.5	6.6	541	2 B75332	transporter, proba
24	142	5.0	544	2 S75696	melibiose carrier
25	139.5	4.9	454	2 A75444	hypothetical prote
26	138.5	4.8	430	2 B75217	transporter PAB217
27	136	4.8	418	2 B87536	membrane protein,
28	133	4.6	389	2 G83413	probable MFS trans
29	128.5	4.5	422	2 G83503	probable MFS trans

ALIGNMENTS

RESULT 1

T14340

sucrose-proton transport protein - carrot

N;Alternate names: sucrose/H+ symporter protein

C;Species: Daucus carota (carrot)

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000

C;Accession: T14340

R;Shakya, R.; Sturm, A.

Plant Physiol. 118, 1473-1480, 1998

A;Title: Characterization of source- and sink-specific sucrose/H+ symporters from carrot

A;Reference number: 217991; MUID:99063785; PMID:9847123

A;Accession: T14340

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-515 <SHA>

A;Cross-references: EMBL:Y16768; NID:g2969883; PIDN:CAA76369.1; PID:g2969884

A;Experimental source: cultivar Namtalse; root

C;Genetics:

A;Note: SUT2

C;Superfamily: common tobacco sucrose transport protein

Query Match	12.1%	Score 347.5	DB 2	Length 515
Best Local Similarity	25.9%	Pred. No. 1e-18		
Matches	131	Conservative	80	Mismatches 215; Indels 79; Gaps 17;
Qy	17	QLLLNLLTFGLVCLAAAGITYVPLLEVGVEKFMVTLGIPVLGLVCPVLLGSASD	76	
Db	34	KLVLVAATAAGVQFGWALQLSLTPYVQLLIPHKWAAIWLCPGISGMLVQPIVGYSD	93	
Qy	77	HWRGRCRRPPFIWALSIGILLSLFLIPRAGWLAGL-----LCPDPRPLEALLILGV	129	
Db	94	HCQSSFGRRPPFTASGAGCVAISVILI---GFAADISYKAGDDMSKTLKPRAVTVFVIGF	150	
Qy	130	GLIDFCQCVQCTPLEALLSLDFR-DPDHCRQAVSVYAFMISLGGCLGV-----LL	178	
Db	151	WILDVANNMLQGPCRALLADLCSDTRRMSANAFYFFWAVGNILGYAAGSYNNLYKLF	210	
Qy	179	PAIDWDTSAIAPYLGTOECLFGLLTLIFTCVAATLLVAEEAALGTPTEAGISAPSL	238	
Db	211	PFS--KTHACDLYCANLKSCFIISALLIITVVALSVVRENS--GPPDDADAABEP---	263	
Qy	239	PHCCPCBARLAFNLGALLPRLHQLCCRMRTLRRLFVAELCSWMLMTFLFTDFVGE	298	
Db	264	----PSSGKIPV--FGELLGALKDL----PRPMLLLIVTCLNIAWFPFLFDTDMGR	313	
Qy	299	GLYGVPRABPGTEARHYDEGVRMGSLGLFLOCAISLVFSLVMDRLVQRFGRVAVLAS	358	
Db	314	EIVGGT--AQGG-----KLYDQGVRAAGALGLLLSVLGLTSIAVEYLVRGVGVKI-LWG	366	
Qy	359	VAAFPVAAAGATCL-----SHSV-----AVVTASA-----ALTGFTFSALQIL	395	
Db	367	FVNFIILAIGLVMTVVVSVKVAHQHRSAGVQGLPPSAGVKAGALSLFSILGIPLSITYSI	426	

C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C;Accession: G84441
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Unayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: G84441
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-594 <STO>
A;Cross-references: GB:AE002093; NID:g3461813; PIDN:AAC32907.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g02860
A;Map position: 2
C;Superfamily: common tobacco sucrose transport protein

Query Match 11.5%; Score 328.5; DB 2; Length 594;
Best Local Similarity 24.1%; Pred. No. 3.3e-17;
Matches 120; Conservative 72; Mismatches 174; Indels 131; Gaps 18;

Qy 22 NLLTFGLVCLAAAGTY-----VPLLLEVGVEEKFMTWVLGIGPVGLGVCVPLGSG 73
Db SLVTLVLSCTVAAGVQFGWALQSLTPTVQTLGSHAFSPFWLGGTGLWQVPGI 118
Qy 74 ASDHWRGRRRRPFIWALSIGLISLFLIPRAGWLA--GLLCPD-----PRPLE 121
Db WSDKCTSKYGRRRPFIWALSIGLISLFLIPRAGWLA--GLLCPD-----PRPLE 175
Qy 122 LALLILGVLLDFCGOVCTPFLAALLSDFRDPHCRQAVSYAFWISLGCGLGVLPAI 181
Db AVFTIIGFWLLDANNTVQGPALLADL--SGPDQRNTANAVFCLMWAIGNILGFSAGS 234
Qy 182 ----DWDTSALAPYLGTQECFLG--LTLPLT-CVAATLLVAEEAALGPTPE 228
Db GKWOEW-----FPFL-TSRACCAAGNLKAAFLAVALVELTICLTVTYFAKEIPTSNKP 288
Qy 229 AE-GLSAPLSPHCCPCARLAFLNL-----LQSGLEHSLKNGTANGIKYVERVDTDFGENSEHQ 340
Db TRIQDSAPLLDD-----GALLPRLHQLCCMRPTRLRLFAELCSMMALMTFTLFTDFVGEGLYQGV 305
Qy 341 DETVVDGPGSVLVNLTSLRHPAMHSVLIVMALTWLSWFFFLFDDTDWNGREYVHCDP 400
Qy 306 RAEPCTEARRHYDGVNMGSLGLFLQCAISLVFSLVMDRLVQRTTRAVY--LASVAAPPV 364
Db TGD--SLHMLYDQGVREGALGLLNSVVLGSISSFLIEPQMCRMGARVWALSNTVPAC 458
Qy 365 AAGATCLS-----HSVAVVTAASALTGTFPSALQILPYTLASLY 403
Db MAGTAVISLMSLSDKNGIEYMRGNETTRAAVIV--FALLGPPLAITYSVPPSVTA-- 514
Qy 404 HREKQVFLPKYRGDTGG 420
Db -----EVTADSGG 522

RESULT 5
S28052
sucrose transport protein - spinach
N;Alternate names: sucrose carrier protein; sucrose permease
C;Species: Spinacia oleracea (spinach)
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 21-Jul-2000
C;Accession: S28052
R;Riesmeier, J.W.; Wallmitzer, L.; Frommer, W.B.
EMBO J. 11, 4705-4713, 1992
A;Title: Isolation and characterization of a sucrose carrier cDNA from spinach by functi
A;Reference number: S28052; MUID:93099843; PMID:1464305
A;Accession: S28052
A;Molecule type: mRNA
A;Residues: 1-525 <RIE>

A;Cross-references: EMBL:X67125; NID:g21318; PIDN:CAA47604.1; PID:g21319
C;Superfamily: common tobacco sucrose transport protein
C;Keywords: transmembrane protein

Query Match 11.3%; Score 324.5; DB 2; Length 525;
Best Local Similarity 23.2%; Pred. No. 5.8e-17;
Matches 129; Conservative 91; Mismatches 196; Indels 141; Gaps 20;

Qy 23 LLTFGLVCLAAAGTY-----VPLLLEVGVEEKFMTWVLGIGPVGLGVCVPLGSGA 74
Db LKGLGLVASAAGVQFGWALQSLTPTVQTLGIPHTWAAVYIWLGGPISGMIVQVLGVY 94
Qy 75 SDHWRGRRRRPFIWALSIGLISLFLIPRAGWLA--GLLCPD-----RPLELALIL 127
Db SDRCTSRFRRRPFIWALSIGLISLFLIPRAGWLA--GLLCPD-----RPLELALIL 151
Qy 128 GVGLLDFCGOVCTPFLAALLSDFR--DPDHCROAVSYAFWISLGCGLGVLPAID---- 182
Db GFWILDVANNTLQGPCRALLADMAAGSQTTRYANAFSFFMALGNITGVAAGSYRLYT 211
Qy 183 ----WDSALAPYLGTQECFLG--LTLPLT-CVAATLLVAEEAALGPTPEAELGSLAPS 237
Db VFPFTKTAACDVYCANLKSFFISITLLIVLITLALSVMKERTIDIEQEEEDLKRN 271
Qy 238 SPHCCPCARLAFLNGLLPRHLQCCMRPTRLRLFAELCSMMALMTFTLFTDFVG 297
Db SSGC----ARLPF--FQQLIGALKDL-----PKMLILLLVLTALNWIAMPFLLFDTDMWG 321
Qy 298 EGLYQGVPRAPGPTTEARRHYDGVNMGSLGLFLQCAISLVFSLVMDRLVQRTTRAVY 355
Db KEVYGGT-----VGEKLYDQGVHAGALGMINSVLGVNMSLIEGLARVMVGAKRLWG 375
Qy 356 LASVAAPFVAAGATCLSHSVAVVTAASALTGTFPSALQILPYTLASLYHREKQVFLPKY 415
Db IVNILL-----AVCLANTV-LVTKSAB-----HFRDSHIM----- 405
Qy 416 GDTGASSEDLSMTSFLPGPKGAPFPNGHVAGSGGLPPPPALCGASACDVSVRVVVG 475
Db -----GSAVPPPPPA--GVKGGALAIFAVLG 429
Qy 476 EPTEARV-VP-----GRGICLDLAILDSAPLLSQV-----AP 506
Db IPLAITSIFPALASIFSASSGQGLSLGVLNLAIVVPMQFVSVTSVPDMFEGGNLP 489
Qy 507 SLFMGSIVQLSQSVTAY 523
Db AFVVGAAVATASAVLSP 506

RESULT 6
S38196
sucrose transport protein SUC2 - Arabidopsis thaliana
N;Alternate names: sucrose-proton symporter SUC2
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 01-Dec-2000
C;Accession: S38196; T00773
R;Sauer, N.K.
submitted to the EMBL Data Library, October 1993
A;Description: SUC1 and SUC2: two sucrose transporters from Arabidopsis thaliana.
A;Reference number: S38196
A;Accession: S38196
A;Molecule type: mRNA
A;Residues: 1-512 <SAU>
A;Cross-references: EMBL:X75382; NID:g407091; PID:g407092
R;Vysotskaia, V.S.; Schwartz, J.R.; Toriumi, M.; Yu, G.; Oji, O.; Kwan, A.; Liu, S.; Li,
rtz, D.; Li, Y.; Palm, C.J.; Shinn, P.; Sun, H.; Davis, R.W.; Ecker, J.R.; Federspiel, N
submitted to the EMBL Data Library, July 1998
A;Description: Arabidopsis thaliana chromosome 1 BAC T22J18 sequence, complete sequence.
A;Reference number: Z14202
A;Accession: T00773
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-490, 'A', 492-512 <VYS>

A;Cross-references: EMBL:AC003979; NID:g3172156; GSPDB:GN00059; ATSP:T22J11
A;Experimental source: cultivar Columbia
C;Genetics:
A;Gene: SUC2; ATSP:T22J18.12
A;Map position: 1
A;Introns: 419/3; 441/1; 455/3
C;Superfamily: common tobacco sucrose transport protein
C;Keywords: sugar transport

Query Match 11.3%; Score 323.5; DB 2; Length 512;
Best Local Similarity 24.1%; Pred. No. 6.7e-17;
Matches 123; Conservative 92; Mismatches 195; Indels 101; Gaps 18;

QY 14 RKQALLVNLTFGLVCLAAAGITYVPPPLLEVGVEEKFMTVLGIGPVGLVCVPLGSG 73
DB 28 LRKIIISVSSIAAGVQFGWALQSLTTPYVQLLGIHPKWSLIWLCGPISGMLVQPIVGY 87
QY 74 ASDHWRGRRRRPFVWALSGLLSLFLIPRAGWLAGL-----GYAADIGHSMGDQDLPKPKTRAIIFA 126
DB 88 HSDRCTSRFRGRRPFVIVAGLVTVAVFLI---GYAADIGHSMGDQDLPKPKTRAIIFA 144
QY 127 LGVGLLDFCGQVCTPLEALLSDLFR-DPDHCRQAVSVYAFMISLGGCLGY-----LL 178
DB 145 LGFWILDVANNTLQGPCRAFLADLSAGNAKKTRTANAFFSFFMAVGNVGLVGAAGSYRNLY 204
QY 179 PAIDWD-TSALAPYLGTQEBECLFGLTLITLTCVAATLLVAEEAALGPTPAEGLSAPSL 237
DB 205 KVPFMTESCDLYCANLKTCCFLSITLL-LIVTFVSLCYVKEKFWPTPEPTADGKA----- 259
QY 238 SPHCCPCRARLAFRNIGALLPRHLQCCRPRTLRLRFLVAELCSWMALMTFTFLFYTDVFG 297
DB 260 -----SNVPP--FGEIFGAFKEL-----KRPMMMLLIIVTALNWIWFFLLFDTDMWG 305
QY 298 EGYQGVPRAPGTEARRHYDEGVRMGSGLFLQCAISLVFSLVMDRLVORFGTRAVYLA 357
DB 306 REVYGNDSATATAASKLYNDGVYRAGALGLMLNAIVLGFMSLGVVEIGRKLIG-CAKRLW 364
QY 358 SVAAPVVAAGATCLSHSVAV-----VTASA-----ALTG-----FT 388
DB 365 GIVNFIL---AICLAMTVVTVTKQENHRRDHGAKTGGPNNVTAGALTFLAILGIPQAIT 421
QY 389 FSALQILPYTLASLYHREK-----OVFLPKYRGDTGCGASSEDLSMTSFLPGP 435
DB 422 FS-----IPFALASIFSTNSGAGQGLSLGVNLAIIVPQMVISVGGGFDELFGGNGNIPA- 476
QY 436 KPGAPFNHGVAGGSGLL-----PPPPA 459
DB 477 -----FVLGAIAAAVSGVLALTLPSPPPDA 502

RESULT 7
G86360
Probable sucrose-proton symporter SUC2 protein T22J18.12 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Nov-2001
C;Accession: G86360
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiet, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; PMID:21016719; PMID:11130712
A;Accession: G86360
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-512 <STO>
A;Cross-references: GB:AE005172; NID:g3287687; PIDN:AAC25515.1; GSPDB:GN00141
C;Genetics:

A;Map position: 1
C;Superfamily: common tobacco sucrose transport protein

Query Match 11.3%; Score 323.5; DB 2; Length 512;
Best Local Similarity 24.1%; Pred. No. 6.7e-17;
Matches 123; Conservative 92; Mismatches 195; Indels 101; Gaps 18;

QY 14 RKQALLVNLTFGLVCLAAAGITYVPPPLLEVGVEEKFMTVLGIGPVGLVCVPLGSG 73
DB 28 LRKIIISVSSIAAGVQFGWALQSLTTPYVQLLGIHPKWSLIWLCGPISGMLVQPIVGY 87
QY 74 ASDHWRGRRRRPFVWALSGLLSLFLIPRAGWLAGL-----GYAADIGHSMGDQDLPKPKTRAIIFA 126
DB 88 HSDRCTSRFRGRRPFVIVAGLVTVAVFLI---GYAADIGHSMGDQDLPKPKTRAIIFA 144
QY 127 LGVGLLDFCGQVCTPLEALLSDLFR-DPDHCRQAVSVYAFMISLGGCLGY-----LL 178
DB 145 LGFWILDVANNTLQGPCRAFLADLSAGNAKKTRTANAFFSFFMAVGNVGLVGAAGSYRNLY 204
QY 179 PAIDWD-TSALAPYLGTQEBECLFGLTLITLTCVAATLLVAEEAALGPTPAEGLSAPSL 237
DB 205 KVPFMTESCDLYCANLKTCCFLSITLL-LIVTFVSLCYVKEKFWPTPEPTADGKA----- 259
QY 238 SPHCCPCRARLAFRNIGALLPRHLQCCRPRTLRLRFLVAELCSWMALMTFTFLFYTDVFG 297
DB 260 -----SNVPP--FGEIFGAFKEL-----KRPMMMLLIIVTALNWIWFFLLFDTDMWG 305
QY 298 EGYQGVPRAPGTEARRHYDEGVRMGSGLFLQCAISLVFSLVMDRLVORFGTRAVYLA 357
DB 306 REVYGNDSATATAASKLYNDGVYRAGALGLMLNAIVLGFMSLGVVEIGRKLIG-CAKRLW 364
QY 358 SVAAPVVAAGATCLSHSVAV-----VTASA-----ALTG-----FT 388
DB 365 GIVNFIL---AICLAMTVVTVTKQENHRRDHGAKTGGPNNVTAGALTFLAILGIPQAIT 421
QY 389 FSALQILPYTLASLYHREK-----OVFLPKYRGDTGCGASSEDLSMTSFLPGP 435
DB 422 FS-----IPFALASIFSTNSGAGQGLSLGVNLAIIVPQMVISVGGGFDELFGGNGNIPA- 476
QY 436 KPGAPFNHGVAGGSGLL-----PPPPA 459
DB 477 -----FVLGAIAAAVSGVLALTLPSPPPDA 502

RESULT 8
S48789
sucrose transport protein - common tobacco
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 26-May-2000
C;Accession: S48789
R;Buerkle, X.Y.Z.; Frommer, W.B.
submitted to the EMBL Data Library, October 1994
A;Reference number: S48789
A;Accession: S48789
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-507 <BOE>
A;Cross-references: EMBL:X82276; NID:g575350; PID:g575351
C;Superfamily: common tobacco sucrose transport protein

Query Match 11.1%; Score 317.5; DB 2; Length 507;
Best Local Similarity 24.3%; Pred. No. 1.9e-16;
Matches 125; Conservative 87; Mismatches 203; Indels 99; Gaps 18;

QY 4 RLWVRLRRHRKQALLVNLTFGLVCLAAAGITYVPPPLLEVGVEEKFMTVLGIGPVGL 63
DB 23 KLV-----KIIVASIAAGVQFGWALQSLTTPYVQLLGIHPKFSFVWLCGPIS 72
QY 64 GLVCPVLLGSASHWRGRRYORRRPFWALSGLLSLFLIPRAGWLAGLCPDP-----R 118
DB 73 GMIVQPVVGYSDNSCSRRFRGRRFTAAAGAAVTVIAVFLIGFAADL-CHATGDPGLGKSK 131
QY 119 PLELALLVGLGLDFCGQVCTPLEALLSDLPRDPDHCRQAVSVYAFMISLGGCLGYLL 178

```
Db 132 PRATAVVGVFWLVDVANNMLOQPCRALLDLGGKARMTSNAPFFSFFMAVGNVGLYAA 191
Qy 179 PAID-----WDTALAPYLGTECELF-GLLTLELTCTVAATLLVAEEAALGPTSPA 229
Db 192 GSYRLCKIPFSTKTPACDIICANLKSCFFTAFLVLLSLTILALT--VRENEPEKDEH 249
Qy 230 EGLSAPSLPHCCPCARLAFRNIGALLPRHLQCCMRPTLRLLFVAELCSWMLMTFT 289
Db 250 E-----IDKAGARKSKVPF--FGEIFGALKDL-----PRPMWILLVTSLNWIARPPFF 297
Qy 290 LFYTFDVEGELYOGVPRAPGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVORF 349
Db 298 LYDTDMWAKYVGG--KVGDG---RLYDLGVHAGALGLLNSVVLGFMSSLSVFLGKKI 351
Qy 350 GTRAVYLASVAAPFVAAGATCLSHSVAVV-----TASAALTGTFTFSALQI-- 394
Db 352 G-GVKRLWGLNFVL---AVCMALTVLVTKVAEKSRQVDAHGTLMAPTSGVKGICALTLFA 407
Qy 395 -----LPYTLASLYHREK-----QVFLPKYRGDTGGASSEDLSMTS 430
Db 408 VLGIPLAVTFSPFALASIFSSNAGSGQGLSLGVNLAIIVVQMLVSIAGGPDLDLFGG 467
Qy 431 FLPGPKPCAPFNGHVAGGSG-----LLPPTPA 459
Db 468 NLPG-----FIVGAVAAASGILALTMLPSPPA 495

RESULT 9
S43142
sucrose transport protein - castor bean
N:Alternate names: sucrose carrier
C:Species: Ricinus communis (castor bean)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 26-May-2000
C:Accession: S43142
R:Weig, A.; Komor, E.
submitted to the EMBL Data Library, March 1994
A:Description: A sucrose carrier from Ricinus communis.
A:Accession: S43142
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-533 <WEI>
A:Cross-references: EMBL:Z31561; NID:G468561; PID:G468562
C:Superfamily: common tobacco sucrose transport protein

Query Match 10.9%; Score 311; DB 2; Length 533;
Best Local Similarity 23.3%; Pred. No. 6.2e-16;
Matches 134; Conservative 95; Mismatches 225; Indels 120; Gaps 19;

Qy 17 QLLVNLTLFGLVCLAAAGITVPPLLLEVGVEEKFTMTVLGIGPVGLVCVPLLGASD 76
Db 37 KVMVASTAAGIQFGWALQSLTTPYVQLLGIPTHTWAAFIWLCGPIGMLVQPIVGHSD 96
Qy 77 HWRGRRRRPFIWALSIGILLFLIPRA---GWLAG-LLCPDPRPLEALLILGVGLL 132
Db 97 RCTSRFGRRRPIASGAFAVAIFVLCYADLGHLSGDSLKSPKTRAIAlFVVGWIL 156
Qy 133 DFCGVCTPTEALLSDLF-RDPDHCROAVSVYAFMISLGGCLGYLLPAID-----W 183
Db 157 DVANNMLOQPCRALLDLSGTSQKTRTANALFSPFMAVGNVGLVGAAYTHLYKLPFT 216
Qy 184 DTSALAPVLTQBECLFGLLTLLFLTCVAATLLVAEEAALGPTSPAEL-----SAP 235
Db 217 KITACDVYCANLKCFFISIVLLLSLTVALSYVREK-----FWSPDQAVDVAEDDTASQA 272
Qy 236 SLSPHCCPCARL--AFRNIGALLPRHLQCCMRPTLRLLFVAELCSWMLMTFTLYT 293
Db 273 SSSAQPMFPFGEILGAFNL-----KRPWILLVLTCLNWIAPFLLPDT 318
Qy 294 DFCVGEGLYOGVPRAPGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVORF 353
Db 319 DMGREVYGG--DSSGSAEQKLVDYRGVVRAGALGLMLNSVVLGFTSLGVEVLARGVG-GV 375
```

```
Qy 354 VYLASVAAPFVAAGATCLSHSVAVVVTASAALTGTFTFSALQILPYTLASLYHREKQVFLPK 413
Db 376 KRLWGIWNFVL---AVCLAMTV-LVTQABST-----RR 405
Qy 414 YRGDTGASSEDLSMTSFLPCPKPGAPFNGHVAGGSGLLPPPPALCGASACDVSVRVV 473
Db 406 FATVSGGAK-----VP-----LPPP---SGVKAGALALFAV 433
Qy 474 VGEPTEARV-VP-----GRGICLDLAILDSQFAPSLFMSISIVQL--SQS 519
Db 434 MGVPQAITYPFALASIFPNTSGAGGGLSLGVNLISIVIPOMIVSVAAGPMDALFGGN 493
Qy 520 VTAYMSAAGLGLVAIYFATQVVDKSDLAKYSA 553
Db 494 LPAFVVGAVAAASGIFALTMLPSPQPDMPSAKA 527

RESULT 10
TI4339
sucrose-proton transport protein - carrot
N:Alternate names: sucrose/H+ symporter protein
C:Species: Daucus carota (carrot)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession: TI4339
R:Shakya, R.; Sturm, A.
Plant Physiol. 118, 1473-1480, 1998
A:Title: Characterization of source- and sink-specific sucrose/H+ symporters from carrot
A:Reference number: 217991; MUID:99063785; PMID:9847123
A:Accession: TI4339
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-501 <SHA>
A:Cross-references: EMBL:Y16766; NID:G2969886; PIDN:CAA76367.1; PID:G2969887
A:Experimental source: cultivar Namtase; leaf
C:Genetics:
A:Gene: SUTra
C:Superfamily: common tobacco sucrose transport protein

Query Match 10.7%; Score 305.5; DB 2; Length 501;
Best Local Similarity 21.6%; Pred. No. 1.5e-15;
Matches 124; Conservative 91; Mismatches 213; Indels 145; Gaps 14;

Qy 18 LLLVNLTLFGLVCLAAAGITVPPLLLEVGVEEKFTMTVLGIGPVGLVCVPLLGASDH 77
Db 30 LLRVASVACGIQFGWALQSLTTPYVQELGIPHAWSSIIWLCGLSLGULLVQPIVGHMSDQ 89
Qy 78 WRGYRRRRPFIWALSIGILLFLIPRAGWLAGLL--CPDPRPLEALLILGVGLLDQFC 135
Db 90 CTSKYGRRRPFIVAGGTAAIILAVIIIAHSADIGLLGDTADNKTMAIVAFVIGFWILDVA 149
Qy 136 GQVCFTPLEALLSDLF-RDPDHCROAVSVYAFMISLGGCLGYLLPAIDMDTSALAPYLG 194
Db 150 NNMTQGPCRALLDLADTNDARRTRVANAYFSLFMAIGNVLGY-----ATGASGV 199
Qy 195 QEELCFGLLTLLFLTC-----VAATLLVAEEAAL-----GPTPEAEL 232
Db 200 KYVFPFSLTSCINCANLKSAPYDIIIFIITTYISIAKERPRISSQDGPQFSEDGT 259
Qy 233 SAPSLPHCCPCARLAFRNIGALLPRHLQCCMRPTLRLLFVAELCSWMLMTFTLY 292
Db 260 AQ---SGHI-----EAFWELFGTFRLLPGSVVWILLVTCNMGWPPFILD 305
Qy 293 TDFVGEGLYOGVPRAPGTEARRHYDEGRVMSGLFLQCAISLVFSLVMDRLVORGTR 352
Db 306 TDMGRIYGGP-----NOGOSYDGVRMGAPFGLMMNSVVLGITSVLMEKLCRIWSSG 359
Qy 353 AVYLASVAAPFVAAGATCLSHSVAVVVTASAALTGTFTFSALQILPYTLASLYHREKQVFLP 412
Db 360 FMW-----GLSNILMTCFFAMLLITFIKNDY----- 388
Qy 413 KYRGDTGASSEDLSMTSFLPGPKPGAPFNGHVAGGSGLLPPPPALCGASACDVSVRV 472
```


Db 38 KLALVASTAAGVQFGWALQSLTTPVQLLGPHTWAPYIMLCGPISGMVQPTVGYTSD 97
Qy 77 HWRGRRRRPPIWALSIGLISLFLIPRAGWLA--GLLCPD-----PRPLELALLIGV 129
Db 98 RCTSKGRRSPFI--AVGATLVGFAVSLIGFAADIGHATGDPNGVFKPRAIAVFFVGF 154
Qy 130 GLLDFCGQVCFPTLEALLSDLFRD--HCRQAYSVYAFMISLGGCLGY-----LLP 179
Db 155 WILDVANNTLQGPCRALLADMAAGSQAKTRYANAFSPFPMALGNIGGYAGSYGRLYTVFP 214
Qy 180 AIDWNTSALAPYLGTQBECLFGLLTLILFUTCVAAITLLVAEEAALGPTPEAGLSAPSLSP 239
Db 215 FT--HTKACDITYCANLKSCFFISITLLIIVLTILALSVVRER----PFTLDEIQEENLKN 268
Qy 240 HCCPCRARLARNLGCALLPRHLQCCRMPTLRLRFLVAELCSWMLMFTLFTYDTPVCEG 299
Db 269 NTGGC--ARLPP--FQQLFGALKDL-----PKMLIILLLVTLNWIAMFPLFDPTDWMGKE 321
Qy 300 LYQGVPRABPGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVORFG--TRAVYLA 357
Db 322 VYGGT-----VGECKAYDMGVHAGALGLMINSVVLGINSGLIEKLARLVGGVKELGWIV 375
Qy 358 SVAAPFVAAGATCLSHSVAVVTA-----ALTFGP 387
Db 376 NLIL-----AVCLAMTI--LVTKSAHYRATHVPCGAIPLPPPPGVKGGALAIFAVLGI 428
Qy 388 FPSALQILPYTLASLYHREK-----QVFLPKYRGDTGGASSEDLSMT 429
Db 429 PLAITFSPFALASIFSSASSGGQSLGLVNLAIWPFQMFVSVTSGPW-----DAL-- 480
Qy 430 SPLPKPKGAPPNGHVGAGSG-----LLPPPP 458
Db 481 -FGGNLPA--FVGVAVATAAISLFTLLPPPP 511

RESULT 14
F96741
probable sucrose transport protein F17M19.4 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: F96741
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: F96741
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-512 <STO>
A:Cross-references: GB:AE005173; NID:96978914; PIDN:AAF34306.1; GSPDB:GN00141
C:Genetics:
A:Gene: F17M19.4
A:Map position: 1
C:Superfamily: common tobacco sucrose transport protein

Query Match 10.2%; Score 292.5; DB 2; Length 512;
Best Local Similarity 23.8%; Pred. No. 1.5e-14;
Matches 120; Conservative 87; Mismatches 21; Indels 87; Gaps 18;

Qy 17 QLLVNLTLFGLEVCLAGITVYPPLLEVGVEKEKFTMVGLGIPVGLVCVPLGASD 76
Db 33 KIISVASAAGVQFGWALQSLTTPYQLLGPHTWAPYIMLCGPISGMVQPTVGYTSD 92
Qy 77 HWRGRRRRPPIWALSIGLISLFLIPRAGWLA--GLLCPD-----PRPLELALLIGV 129

Db 93 RCEGRGRRRPFPIAGVALVAVSVFLI---GFAADMGHSGFGDKLENKVRTRAIIFLTGTF 149
Qy 130 GLLDFCGQVCFPTLEALLSDLFR--DPDHCRQAYSVYAFMISLGGCLGY-----LL 178
Db 150 WFLDVANNTLQGPCRAFLADLAAGDAKTRVANACFSFFMAVGNVLYGAAGSYTNLHKMF 209
Qy 179 PAIDWNTSALAPYLGTQBECLFGLLTLILFUTCVAAITLLVAEEAALGPTPEAGLSAPSLSP 238
Db 210 PFT--MTKACDITYCANLKSCFFISITLLIIVTSSLSLVYKDK---QWSPQGDKEEKT 263
Qy 239 PHCCPCRARLARNLGCALLPRHLQCCRMPTLRLRFLVAELCSWMLMFTLFTYDTPVCEG 298
Db 264 -----SLFF--FGEIFGAVR---HMKRPMVMLLIIVTINIAMFPLFTYDTPDMMGR 309
Qy 299 GLYQGVPRABPGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMDRLVORFGTRAVYLA 358
Db 310 EYVGG--NSDGDERSKLYDOGVAGALGLMINSILGLFVSLGVESGRKMG--GAKRLWG 366
Qy 359 VAAPFVAAGATCLSHSVAVVTA-----SAALTGFTTFSALQIL----- 395
Db 367 CVNFILAI--LMTVLVTKSAHREIAGPLAGSSGKAGVFSLFTVLGILPLAITYS 423
Qy 396 -PYTLASLYHREKQVFLPKYRGDTGGASSEDLSMTSFLPGKP-----GAPPN---GHVG 447
Db 424 IPFALASIFSTNSGAGQSLGLVNLAIACIPQMTVSVFSSGPLDAQFGCGNLPSFVVGAI 483
Qy 448 AGSGSL-----PPP--PALCGA 463
Db 484 AAVSGVLAITVLRPFPDAPAMSGA 508

RESULT 15
T02982
probable sucrose transport protein - rice
C:Species: Oryza sativa (rice)
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 21-Jul-2000
C:Accession: T02982
R:Hirose, T.; Imaizumi, N.; Scofield, G.N.; Furbank, R.T.; Ohsugi, R.
Plant Cell Physiol. 38, 1389-1396, 1997
A:Title: cDNA cloning and tissue specific expression of a gene for sucrose transporter f
A:Reference number: 214809; MUID:98182940; PMID:9522469
A:Accession: T02982
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-537 <HR>
A:Cross-references: EMBL:D87819; NID:g2723470; PIDN:BAA24071.1; PID:g2723471
A:Experimental source: cultivar Nipponbare, leaf
C:Genetics:
A:Gene: SUT1
C:Superfamily: common tobacco sucrose transport protein

Query Match 10.2%; Score 292; DB 2; Length 537;
Best Local Similarity 26.7%; Pred. No. 1.7e-14;
Matches 115; Conservative 72; Mismatches 167; Indels 76; Gaps 19;

Qy 17 QLLVNLTLFGLEVCLAGITVYPPLLEVGVEKEKFTMVGLGIPVGLVCVPLGASD 76
Db 51 RLILSGMVAGGVQFGWALQSLTTPYQLLGPHTWAPYIMLCGPISGMVQPTVGYTSD 110
Qy 77 HWRGRRRRPPIWALSIGLISLFLI---PRAGWLAAGLCPD-----PRPLELALLIL 127
Db 111 RCTSKGRRRPIYILTGCVLICLAVVIGFSDADIGYANGDYKEDCSVYHGSRWHAIVVVL 170
Qy 128 GVLGLDFCGQVCFPTLEALLSDLF--RDPDHCRQAYSVYAFMISLGGCLGYLLPAID--- 182
Db 171 GFWLDFSNNTVQGPAPALMADLSGRHGP---TANSIFCSWMAGNIGLYSGSGSTNNWH 227
Qy 183 -WNTSALAPYLGTQBECLFGLLTLILFUTCVAAITLLVAEEAALGPTPEAGLSAPSLSP 226
Db 228 KW-----PFLKTRACCEACANLKGAFLVAVIFLSCLVITLIFAKEVPPFKGNAAL--PTK 281
Qy 227 --BPAEGLSAPSLSPHCCPCRARLARNLGCALLPRHLQCCRMPTLRLRFLVAELCSWMA 284

```
Db 282 SNEPAEPEGTC-----PLAVLKGFRNLPTGMPSV-----LIVTCL-TWLS 320
Qy 285 LMTFTLFYDFVGEGLYQGVPR-AEPGTEARRHYDEGVRMGSLGLFLQCAISLVFSLVMD 343
Db 321 WFPFILDYTDWMGRIYHGDPKGTDPOIEA---FNQGVRAAGFGLLINSIVLGFSSFLIE 377
Qy 344 RLQVRFGTTRAVYLASVAAPVVAAGATCL-----SHSVA--VVTTASAALTGFTFSALQ 393
Db 378 PMCRKVGPRVWVWTSNFLVLCIAMAAATALISFWSLKDPHGTVQKAITADKSIKAVCLVIFA 437
Qy 394 ILPYTLASLY 403
Db 438 FLGVPLAVLY 447
```

Search completed: January 25, 2004, 17:05:52
Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2004, 17:01:07 ; Search time 18 Seconds
(without alignments)
1444.765 Million cell updates/sec

Title: US-09-593-793a-113

Perfect score: 2861

Sequence: 1 MVQRLWVSLRLHRKAQLL.....AIYFATQVVPDKSLAKYSA 553

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	489.5	17.1	530	1	MATP_MOUSE
2	488.5	17.1	530	1	MATP_HUMAN
3	324.5	11.3	525	1	STP_SPIOL
4	235.5	8.2	553	1	SUTI_SCHPO
5	142	5.0	544	1	YD74_SYNY3
6	136	4.8	541	1	GT10_HUMAN
7	125.5	4.4	399	1	TCRI_ECOLI
8	121.5	4.2	503	1	PUR8_STRLP
9	119.5	4.2	491	1	AMPG_ECOLI
10	117	4.1	440	1	YHUE_ECOLI
11	117	4.1	495	1	SNVA_SALTY
12	111.5	3.9	640	1	Y051_MYCTU
13	111	3.9	473	1	PHDK_NOCSE
14	110.5	3.9	680	1	CA1A_HUMAN
15	110	3.8	368	1	GALT_HUMAN
16	110	3.8	606	1	NOOC_THETH
17	109.5	3.8	654	1	SPH2_HUMAN
18	109	3.8	477	1	YFUM_RHOCA
19	107	3.7	419	1	CMLA_PSEAE
20	105	3.7	448	1	YJ94_YEAST
21	104	3.6	476	1	YMLB_SALTY
22	103	3.6	465	1	FTSW_MYCLE
23	102.5	3.6	461	1	PUCG_RHOCA
24	102	3.6	438	1	SHIA_ECOLI
25	101.5	3.5	428	1	YXIO_BACSU
26	101.5	3.5	510	1	NANT_YERPE
27	101	3.5	437	1	BRAZ_YEAST
28	101	3.5	471	1	YMLB_PSEAE
29	101	3.5	471	1	YMLB_KLEPN
30	100	3.5	404	1	YAXM_RHISN
31	100	3.5	473	1	YIHO_SALTY
32	99	3.5	386	1	CV03_HUMAN
33	99	3.5	481	1	LMRA_STRLN

ALIGNMENTS

RESULT 1

ID	MATP_MOUSE	STANDARD	PRT	530 AA
AC	P59355			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Membrane-associated transporter protein (AIM-1 protein) (Melanoma antigen AIM1) (Underwhite protein).			
GN	MATP OR AIM1 OR UW.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;			
OX	NCBI_TaxID=10090;			
LN	[1]			
RP	SEQUENCE FROM N.A.			
PC	SPRAIN=ddv; TISSUE=Eye, Kidney, and Uterus;			
RX	MEDLINE=21372467; PubMed=11479596;			
RA	Fukumachi S., Shimada A., Shima A.;			
RT	"Mutations in the gene encoding B, a novel transporter protein, reduce melanin content in medaka."			
RL	Nat. Genet. 28:381-385(2001).			
LN	[2]			
RP	SEQUENCE FROM N.A., AND VARIANTS UW-DBR ASN-153 AND PRO-435.			
RX	MEDLINE=21473748; PubMed=11574907;			
RA	Newton J.M., Cohen-Barak O., Hagiwara N., Gardner J.M., Davissson M.T., King R.A., Brilliant M.H.;			
RT	"Mutations in the human orthologue of the mouse underwhite gene (uw) underlie a new form of oculocutaneous albinism, OCA4."			
RL	Am. J. Hum. Genet. 69:981-988(2001).			
CC	-1- FUNCTION: Melanocyte differentiation antigen. May transport substances required for melanin biosynthesis (By similarity).			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein; melanosome (By similarity).			
CC	-1- TISSUE SPECIFICITY: Melanocytes, eyes, kidney and uterus.			
CC	-1- DISEASE: Defects in MATP are the cause of the UW-dbr phenotype that results in loss of nearly all pigmentation in the homozygous state.			
CC	-1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL; AF360357; AAK81713.1; -			
DR	MGI; MGI:2153040; Matp.			
DR	Melanin biosynthesis; Transmembrane; Antigen; Glycoprotein; Vision; Disease mutation; Albinism.			
KW	DOMAIN 1 45			
FT	DOMAIN 46 66			
FT	TRANSMEM 67 68			
FT	DOMAIN 69 89			
FT	TRANSMEM 69 89			

34	98.5	3.4	613	1	NUOL_ECOLI
35	98.5	3.4	685	1	PHUB_SALTY
36	98	3.4	379	1	CYB_SORCI
37	98	3.4	496	1	GNR3_CHICK
38	97.5	3.4	354	1	RNED_PSEST
39	97.5	3.4	385	1	PI2R_BOVIN
40	97.5	3.4	433	1	DCUA_WOLSU
41	97.5	3.4	1103	1	CYGD_HUMAN
42	97	3.4	336	1	CYB_SORHA
43	97	3.4	336	1	CYB_SORMO
44	97	3.4	336	1	CYB_SORVA
45	97	3.4	370	1	GALT_RAT

P33607	escherichia
O87656	salmonella
O79451	sorex ciner
P28568	gallus gall
O8evn4	pseudomonas
P79393	bos taurus
O34245	wolinnella s
Q02846	homo sapien
O79452	sorex hayde
O79969	sorex monti
O80019	sorex vagra
O88626	rattus norv


```

DR Pfam; PF00083; sugar_tr; 1.
DR TIGRFAMS; TIGR01301; GPH_sucrose; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; FALSE NEG.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; FALSE NEG.
KW Transmembrane; Transport; Sugar transporter; Symport.
FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 38 58 1 (POTENTIAL).
FT TRANSMEM 72 92 2 (POTENTIAL).
FT TRANSMEM 107 127 3 (POTENTIAL).
FT TRANSMEM 145 165 4 (POTENTIAL).
FT TRANSMEM 184 204 5 (POTENTIAL).
FT TRANSMEM 230 250 6 (POTENTIAL).
FT TRANSMEM 295 315 7 (POTENTIAL).
FT TRANSMEM 338 358 8 (POTENTIAL).
FT TRANSMEM 373 393 9 (POTENTIAL).
FT TRANSMEM 422 442 10 (POTENTIAL).
FT TRANSMEM 455 475 11 (POTENTIAL).
FT TRANSMEM 488 508 12 (POTENTIAL).
FT DOMAIN 509 525 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 525 AA; 54992 MW; 018347A4D2CC1CC6 CRC64;

Query Match 11.3%; Score 324.5; DB 1; Length 525;
Best Local Similarity 23.2%; Pred. No. 2.5e-16;
Matches 129; Conservative 91; Mismatches 136; Indels 141; Gaps 20;

Qy 23 LLTFLGVLCUAGITY-----VPLLLEVGVEKFMVVLGIGVGLVGLVPLLSGA 74
Db 35 LKGLGLVASVAQVGFQWALQSLTTPVOLLGIPHTWAAVYVLCGPISGMIVPLVGY 94
Qy 75 SDHWGRYGRRRPTIWLISGILLSLPLIPRAGWLA--GLLCDDP-----RPLELALL 127
Db 95 SDRCTSRFGRRRPTIAGAAALVAVAVGLI---GPAADIGAASGDPTGNVAKPRAIAFV 151
Qy 128 GVGLLDGCGVCFPLEALLSLDFR-DPDHCRQAVSVYVFMISLGGCLGYLLPAID--- 182
Db 152 GFWLIDVANNLTQPCRRALLADMAAGSQTKTRYANAFSFMALNGTGGVAAAGSYRLYT 211
Qy 183 ----WDTSAIAPYLGTOECLFGLLT-LIFELTCAATLLVAEEAALGFTBPAGLSAPSL 237
Db 212 VFPTKTAACDVYCANLKSFFISITLLIVLTALSVKXERQITIDEIQEEDLNKNN 271
Qy 238 SPHCCPCRLAFRNLCALLPRLHQLCRMPRLRLFLVLAELCSWMAIMTFTLEYTDVFG 297
Db 272 SSGC---ARLPF--FQGLIGALKDL----PKMLILLVLTALNWIAMFPFLFDTDWNG 321
Qy 298 EGLVQGVPRAEPTGARRHYDEGVMSGLFLQCAISLVFLVMDRLVORFG--TRAVY 355
Db 322 KEVYGGT-----VGEKLDQGVHAGALGLMINSVLGWSLSIEGLARVGGAKELWG 375
Qy 356 LASVAAPFVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKYR 415
Db 376 IVNIIL-----AVCLAMTV-LVTKSAE-----HFRDSSHIM--- 405
Qy 416 GDTGGSSEDSMTSLFPLPGPKGAPFNHGVAGCGSGLLPPPPALCGASACDVSVRVVVG 475
Db 406 -----GSAVPPPPPPA--GVKGALAIFAVLG 429
Qy 476 EPTEARV-VP-----GRGICLDLAILDASFLLSQV-----AP 506
Db 430 IPLAITSIPPALASIFSASSGSGGLSLGVNLIAIVPQMFVSVTSGPNDAMPGGGNLP 489
Qy 507 SLFMGSIVQLSQSVTAY 523
Db 490 AFVVGAVAATASAVLSF 506

RESULT 4
ID SUT1 SCHPO STANDARD; PRT; 553 AA.
AC O14091; OSUUG60;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)

```

General alpha-glucoside permease.

SUT1 OR SPAC2F3.08.

Schizosaccharomyces pombe (Fission yeast).

Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

Schizosaccharomycetales; Schizosaccharomycetaceae;

Schizosaccharomycetes.

NCBI_TaxID=4896;

[1]

SEQUENCE FROM N.A.

STRAIN=972;

MEDLINE=21848401; PubMed=11859360;

Wood V., Gwilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,

Souros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,

Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,

Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,

Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,

Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,

James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,

Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,

Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,

Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,

Skellon J., Simmonds M., Squares R., Squares S., Stevens K.,

Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,

Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,

Welters I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,

Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,

Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,

Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,

Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,

Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,

Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,

Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,

Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,

Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;

"The genome sequence of Schizosaccharomyces pombe.";

Nature 415:871-880(2002).

[2]

SEQUENCE OF 35-166 FROM N.A.

STRAIN=968 h90;

MEDLINE=20223868; PubMed=10759889;

Ding D.-Q., Tomita Y., Yamamoto A., Chikashige Y., Haraguchi T.,

Hiraoka Y.;

"Large-scale screening of intracellular protein localization in living

fission yeast cells by the use of a GFP-fusion genomic DNA library.";

Genes Cells 5:169-190(2000).

[3]

IDENTIFICATION AND FUNCTION

MEDLINE=21065106; PubMed=11136464;

Reinders A., Ward J.M.;

"Functional characterization of the alpha-glucoside transporter Sut1p

from Schizosaccharomyces pombe, the first fungal homologue of plant

sucrose transporters.";

Mol. Microbiol. 39:445-454(2001).

-!- FUNCTION: Responsible for the transport of maltose and sucrose

into the cell, with the concomitant export of a proton (Symport

system).

-!- PATHWAY: Maltose and sucrose metabolism.

-!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@isb-sib.ch).

EMBL; Z99165; CAB16264.1; --

EMBL; AB027797; BAA87101.1; --

PIR; T38541; T38541.

GeneDB SPombe; SPAC2F3.08; --

Transmembrane; Transport; Sugar transporter; Symport.

FT DOMAIN 1 33 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 34 54 POTENTIAL.
 FT DOMAIN 55 72 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 73 93 POTENTIAL.
 FT DOMAIN 94 111 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 112 132 POTENTIAL.
 FT DOMAIN 133 140 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 141 161 POTENTIAL.
 FT DOMAIN 162 186 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 187 207 POTENTIAL.
 FT DOMAIN 208 216 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 217 237 POTENTIAL.
 FT DOMAIN 238 280 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 281 301 POTENTIAL.
 FT DOMAIN 302 322 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 323 343 POTENTIAL.
 FT DOMAIN 344 424 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 425 445 POTENTIAL.
 FT DOMAIN 446 452 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 453 473 POTENTIAL.
 FT DOMAIN 474 494 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 495 515 POTENTIAL.
 FT DOMAIN 516 521 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 522 542 POTENTIAL.
 FT DOMAIN 543 553 CYTOPLASMIC (POTENTIAL).
 SQ SEQUENCE 553 AA; 61744 MW; 2C31DFC2D88A7C7A CRC64;

Query Match 8.2%; Score 235.5; DB 1; Length 553;
 Best Local Similarity 24.0%; Pred. No. 7.6e-10;
 Matches 82; Conservative 59; Mismatches 153; Indels 47; Gaps 10;

QY 19 LNVLLTGLEVCVLAAGITYVPPILLLEVGVEKFWMTVLGIPVGLVCVPLLGASDHW 78
 DB 39 LTVSLL--GVQLTWSVGLGYTPYLSGLRKEWTSIIWIAGPLTGILIQIAGILSDRV 96
 QY 79 RGRVRRPFTWALSGLILLSLFLIPRAGWAGLLCPD-----PRPLELALLIL 127
 DB 97 NSRIGRRPPFMCASLLGTSLFLM--GW-----APDCLFIFSNVLMKRVTVLTI 148
 QY 128 GVGILLDFCGVCFPLEALLSDLRDPDHCQAYSVYAFMISLGCGLGILPAID-WDTS 186
 DB 149 SIYLLDVAVNVMASTRSLVDSVRS--DQOHEANSWAGRMIGVGNVGLGLYLPLRYIF 207
 QY 187 ALAPYLTQEBCLFGLLFLITCVAATLLVAEEAALGTPRPAEGLSAPSLSPHCCPCA 246
 DB 208 SFNFTQLOVFCVLASISLV-LTVITTFVSR-RFPVEHEKSVAG-----253
 QY 247 RLAFNIGALLPRLHQLCCMRPRLRLFLVAELCSWMALMTFTLYTDFVGEGLYQGVPR 306
 DB 254 -----EIFEFTTMRQSITAPFTLKRICFVQFFAYFGWFFFLFVITYVGLYLRHAPK 308
 QY 307 AEPGTEARRHYDEGVMSGLFLQCAISLVFSLVMDLVQ 347
 DB 309 GH-----EEDWDMATROGSFALLPFAIISLAANTALPLLE 344

RESULT 5

YD74_SNNY3
 ID_YD74_SNNY3 STANDARD; PRT; 544 AA.
 AC P74168;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical symporter sl11374.
 GN SL11374.
 OS Synechocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 OX NCBI_TaxID=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97061201; PubMed=8905231;
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 Miyajima N., Hiroseawa M., Sugiyama M., Sasamoto S., Kimura T.,

RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
 Shimpō S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
 Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 Synechocystis sp. strain PCC6803. II. Sequence determination of the
 entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136(1996).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- SIMILARITY: BELONGS TO THE SODIUM:GALACTOSIDE SYMPORTER FAMILY
 (SGF).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; D90912; BAA18257.1; -;
 DR PIR; S75696; S75696.
 DR InterPro; IPR001927; Na/Gal_symport.
 DR TIGRFAMs; TIGR00792; gph; 1_symport.
 DR PROSITE; PS00872; NA_GALACTOSIDE_SYMPT; FALSE_NEG.
 KW Hypothetical protein; Transport; Transmembrane; Symport;
 Complete proteome.
 FT TRANSMEM 31 51 POTENTIAL.
 FT TRANSMEM 52 72 POTENTIAL.
 FT TRANSMEM 84 104 POTENTIAL.
 FT TRANSMEM 116 136 POTENTIAL.
 FT TRANSMEM 162 182 POTENTIAL.
 FT TRANSMEM 191 211 POTENTIAL.
 FT TRANSMEM 230 250 POTENTIAL.
 FT TRANSMEM 257 277 POTENTIAL.
 FT TRANSMEM 318 338 POTENTIAL.
 FT TRANSMEM 356 376 POTENTIAL.
 FT TRANSMEM 383 403 POTENTIAL.
 FT TRANSMEM 407 427 POTENTIAL.
 FT TRANSMEM 450 470 POTENTIAL.
 FT TRANSMEM 501 521 POTENTIAL.
 SQ SEQUENCE 544 AA; 60064 MW; C91D0EDFF32277EE CRC64;

Query Match 5.0%; Score 142; DB 1; Length 544;
 Best Local Similarity 22.3%; Pred. No. 0.0045;
 Matches 105; Conservative 60; Mismatches 179; Indels 126; Gaps 20;

QY 16 AQLLVNLLFTGLEVCVLAAG-----ITYVPLLELV-GVEKFWMTVLGIPVL 63
 DB 2 SQSLSAEKLHFTTKLAYGAGDFGPAITANILVFLFLTDVAGIPAAAGSVLMIGKIF 61
 QY 64 GLVCVPLLGASDHWRCYGRRRPFIWALSIGILLSLFLIPRAGWAGLLCPDPRPLELA 123
 DB 62 DAINDPIIGLLSDRTSRWGRRLPFW-----LGGMIPFALFYTAQWLIPHSDDRLTNQWG 117
 QY 124 LLILGVGLDFCGVCFPT-----PLEALLSDLRDPDHCQAYSVYAFMIISLGGCGLGVL 179
 DB 118 LFYVVAIA-WAFNLCTYTNVLPYALTPELTQYNE-RTELNSFRAPFSGGSLIL 175
 QY 180 AIDWDTSAAPYLGTQEBCLFGLL-TLIFLTCVAATLLVAEEAALGTPRPAEGLSAPSL 238
 DB 176 YI-----LIAAGLPDRPQQQFGLGVMSVLSISALLWSALRLQEKKEP---ILSPSL- 226
 QY 239 PHCCPCARLA-----FRNLGAL-----256
 DB 227 -----RRRLAPLMAAGITLILAIKSNLGGSGDFYISPFLLILLGWGGFTLR 280
 QY 257 -----LPRHLQCCRPRTLR-LFV--AELCSWMALMTFTLF 291
 DB 281 DSAVEEHLQKLENSPSPGVTEINPLLKQL--KIAFSNRAFLFVIGIYLCSLAVQLTASI 338
 QY 292 YTDVFGEGLYQGVPRAPFGTEARRHYDEGVMSGLFLQCAISLVFSLVMDLVQRFGT 351
 DB 339 LVFVVS--WMGLNEQSGT-----IALAVQ-GTALVLMFVWQALAQFLDK 381

Db 109 GFAISLSMCCIYVSELV-GPRQGVLSVYERAGITVGIILSVNALVALAGTGWGRHM 167
Qy 187 ---ALAPYLGTQEELGGLLIILFUTCAATLLVAEEAALGFTPEAGLSAPSLSPHCCP 243
Db 168 FGWATAP-----AVLOSLSLLFLP--AGT---DETATHKDLPLQGEAPKLGPG--- 212
Qy 244 CRALAFRNLCALLPRLHQLCRMPRTURRFLVAVELCSWMLMTFTLFTYDFVGEGLYQG 303
Db 213 -RPRYSFLDLPRARDNMRG---RTTVGLGLVLFQOLTQCPNVLCVA--STIFFSSVGFHGG 266
Qy 304 VPRAEPTEARRHYDEGVMSGLGFLQALISLVPFVMDLVORFGTRAVYLASVA--- 360
Db 267 -----SSAVLASVGL---GAVKVAATLTMGLVDRAGRALLAGCALMA 308
Qy 361 -----AFFVAAGATCLSHSVAVVTASAALTG----FTFSALQILPYTLASLYHR 405
Db 309 LSVSGIGLVFAVPMDSGSL--AVPNATQGTGLPGDGLQLQDSSLPIPT-----N 360
Qy 406 EKQVFLPKYRGDTGGASEDSLMTSFLPGPKGAPFPNGHVGAGSGLLPPPPALCGASA 465
Db 361 EDQ-----REPLSTAKTKPHPRSGDPSAPRLALSSALPGPP----- 399
Qy 466 CDVSVRVVVGEPTEARVVGPG-----ICLDIALDSAP-----LLSQVAP-- 506
Db 400 -----LPAFGHALLRWTLALLCMVFSAFSGFGPVTWLVLSEIYVE 442
Qy 507 -----SLFMG-----SIVQLQSOTVAYMVSA--GLGLVAIY 536
Db 443 IRGRAFAFCNFNWAANLFLSFLDLGTLGSLWTFLLGLTAVLGLGFIYLP 496

RESULT 7
TCR1_ECOLI STANDARD; PRT; 399 AA.
AC P02982;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tetracycline resistance protein, class A (TETA(A)).
GN TETA.
OS Escherichia coli.
OG Plasmid RPI.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSON=Tn1721;
RX MEDLINE=92192465; PubMed=1312499;
RA Allmeier H., Greenar B., Greck M., Schmitt R.;
RT "Complete nucleotide sequence of Tn1721: gene organization and a
RL novel gene product with features of a chemotaxis protein.";
RN Gene 111:11-20(1992).
RP [2]
RC STRAIN=DH1;
RA Fletcher J.N., Hart C.A., Batt R.M., Saunders J.R.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC PLASMID=RP1; TRANSPOSON=Tn1721;
RX MEDLINE=83299270; PubMed=6310527;
RA Waters S.H., Rogowsky P., Grinstead J., Altenbuchner J., Schmitt R.;
RT "The tetracycline resistance determinants of RPI and Tn1721:
RL nucleotide sequence analysis";
RN Nucleic Acids Res. 11:6089-6105(1983).
RP [4]
RC TOPOLOGY
RX MEDLINE=92388137; PubMed=1517220;
RA Allard J.D., Bertrand K.P.;
RT "Membrane topology of the pBR322 tetracycline resistance protein.
RT TetA-PhoA gene fusions and implications for the mechanism of TetA
RT membrane insertion.";

J. Biol. Chem. 267:17809-17819(1992).
-!- FUNCTION: RESISTANCE TO TETRACYCLINE BY AN ACTIVE TETRACYCLINE
EFFLUX. THIS IS AN ENERGY-DEPENDENT PROCESS THAT DECREASES THE
ACCUMULATION OF THE ANTIBIOTIC IN WHOLE CELLS. THIS PROTEIN
FUNCTIONS AS A METAL-TETRACYCLINE/H+ ANTIporter.
-!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
-!- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
AS THE DRUG RESISTANCE TRANSLOCASE FAMILY).

THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X61367; CAA43643.1; --
EMBL; L29404; AAB83545.1; --
EMBL; X00006; CAA24909.1; --
PIR; A03509; YTECR1.
PIR; J01479; J01479.
InterPro; IPR004734; Drug_resist.
InterPro; IPR007114; MFS.
InterPro; IPR005828; Sub_transporter.
InterPro; IPR005829; Sug_transporter.
InterPro; IPR001958; TCR_Teta.
Pfam; PF00083; sugar_tr_1.
PRINTS; PR01035; TCR_TETA.
TIGRFAMs; TIGR00880; 2_A_01_02; 1.
PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
KW Antibiotic resistance; Transmembrane; Inner membrane; Transport;
KW Antiporter; Ion transport; Hydrogen ion transport; Transposable element;
KW Plasmid.
FT DOMAIN 1 7 CYTOPLASMIC (PROBABLE).
FT TRANSMEM 8 27 PROBABLE.
FT DOMAIN 28 45 PERIPLASMIC (PROBABLE).
FT TRANSMEM 46 66 PROBABLE.
FT DOMAIN 67 79 CYTOPLASMIC (PROBABLE).
FT TRANSMEM 80 100 PROBABLE.
FT DOMAIN 101 103 PERIPLASMIC (PROBABLE).
FT TRANSMEM 104 124 PROBABLE.
FT DOMAIN 125 138 CYTOPLASMIC (PROBABLE).
FT TRANSMEM 139 159 PROBABLE.
FT DOMAIN 160 160 PERIPLASMIC (PROBABLE).
FT TRANSMEM 161 181 PROBABLE.
FT DOMAIN 182 210 CYTOPLASMIC (PROBABLE).
FT TRANSMEM 211 231 PROBABLE.
FT DOMAIN 232 246 PERIPLASMIC (PROBABLE).
FT TRANSMEM 247 267 PROBABLE.
FT DOMAIN 268 277 CYTOPLASMIC (PROBABLE).
FT TRANSMEM 278 298 PROBABLE.
FT DOMAIN 299 299 PERIPLASMIC (PROBABLE).
FT TRANSMEM 300 320 PROBABLE.
FT DOMAIN 321 339 CYTOPLASMIC (PROBABLE).
FT TRANSMEM 340 360 PROBABLE.
FT DOMAIN 361 364 PERIPLASMIC (PROBABLE).
FT TRANSMEM 365 385 PROBABLE.
FT DOMAIN 386 399 CYTOPLASMIC (PROBABLE).
FT CONFLICT 5 5 R -> I (IN REF. 3).
FT CONFLICT 55 55 M -> V (IN REF. 3).
FT CONFLICT 75 75 V -> I (IN REF. 3).
FT CONFLICT 84 84 A -> T (IN REF. 3).
FT CONFLICT 201 203 ASF -> SFV (IN REF. 3).
SQ SEQUENCE 399 AA; 42240 MW; 298427E6B5478374 CRC64;
Query Match 4.4%; Score 125.5; DB 1; Length 399;
Best Local Similarity 22.3%; Pred. No. 0.051;
Matches 104; Conservative 57; Mismatches 136; Indels 169; Gaps 27;
Qy 19 LLVNLTLFGLEVCIAAGITYVPLP---LLEVGVEEKFTMT---LVLGTPVLGVCPVLL 71
Db 7 LIVLSTVALD---AVGILMPVLGRLDLVHSNDVTAHYGILLALYALMQFACAPVL 63

Db 156 MLYSGGLALWL--ADKW-----LQWQ--MYWMAALLIPCIITATILAPE-----PT 198

Qy 227 EPAEGLSAPSLPHCCPCRLARAFRNIGALLPRHLQCCRMPTLRFLFAEL----- 279

Db 199 D-----TIP-----VPKTEQAVVAPLRDFFGRN 222

Qy 280 CSWALMTFTL-----FYTDFVGEGLYQGVPRAPGTEARRHYDEGVMSGLFLQCA 333

Db 223 NAMLILLILLVLYKIGDAFAMSLTTFTLIRGV-----GFDAG---EVGVNKTILGLLATIV 274

Qy 334 ISLVSFLVMDRLVQRFGTRAVLASVAAPFAAGATCLSHSVAVVTASAALTGTFFSALQ 393

Db 275 GALTGGILMQRLSL--FRALLIFGI-----LQGNASAGYW 307

Qy 394 ILPYTLASLHYREKQVLPKYRGDTGGASSEDLSMTSFLPGPKGAPFPNGHVAGGSG 453

Db 308 LLSITDKHLYSMGAAPFFENLCCGMGTSAFVALLMT-----LCNKSFSATQFAL 356

Qy 454 LPPPALCGASACDVSVRVVGEFTEARVPGRGICLDLALDSAFILLSOVA-----PSLFM 510

Db 357 L-----SALSAGRVYVG-VPVAGFVEAHGW-----STFYLFVAAAAPGLIL 398

Qy 511 GSIQLSQ-----SVTAYMVS-----AAGLGLVAIYFATQVVFDKSLAKYS 552

Db 399 LLVCQRTLEYTRVNDNFISRTAYPAGYAFAMTWLAAGVSLVAVLILL-LTWDALDLTHFS 457

RESULT 10

YHJE_ECOLI

ID -YHJE_ECOLI STANDARD; PRT; 440 AA.

AC P37643;

DT 01-OCT-1994 (Rel. 30, Created)

DT 01-OCT-1994 (Rel. 30, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Hypothetical metabolite transport protein yhjE.

GN YHJE OR B3523.

OS Escherichia coli.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OT Enterobacteriaceae; Escherichia.

OX NCBI_TaxID=562;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=K12 / MG1655;

RX MEDLINE=94316500; PubMed=8041620;

RA Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;

RT "Analysis of the Escherichia coli genome. V. DNA sequence of the

RT region from 76.0 to 81.5 minutes."

RL Nucleic Acids Res. 22:2576-2586(1994).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane

CC (Potential).

CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. STRONG, TO

CC H. INFLUENZAE HI0281 AND HI0418.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC ENBL; U00039; AAB18499.1; -.

DR ENBL; AE000429; AAC76548.1; -.

DR PIR; S47743; S47743.

DR EcoGene; EGI2249; yhjE.

DR InterPro; IPR004736; Cit_H_symport.

DR InterPro; IPR007114; MFS.

DR InterPro; IPR005828; Sub transporter.

DR InterPro; IPR005829; Sug transporter.

DR Pfam; PF00083; sugar_tr_1.

DR TIGRFAMs; TIGR00883; 2A0106; 1.

DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.

DR PROSITE; PS00217; SUGAR_TRANSPORT_2; FALSE_NEG.

KW Hypothetical protein; Transport; Transmembrane; Inner membrane;

KW Complete proteome.

FT DOMAIN 1 34 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 35 55 1 (POTENTIAL).

FT DOMAIN 56 66 PERIPLASMIC (POTENTIAL).

FT TRANSMEM 67 87 2 (POTENTIAL).

FT DOMAIN 88 108 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 109 129 3 (POTENTIAL).

FT TRANSMEM 130 150 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 151 167 4 (POTENTIAL).

FT TRANSMEM 168 188 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 189 192 PERIPLASMIC (POTENTIAL).

FT TRANSMEM 193 213 6 (POTENTIAL).

FT DOMAIN 214 248 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 249 269 7 (POTENTIAL).

FT DOMAIN 270 289 PERIPLASMIC (POTENTIAL).

FT TRANSMEM 290 310 8 (POTENTIAL).

FT DOMAIN 311 320 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 321 341 9 (POTENTIAL).

FT DOMAIN 342 345 PERIPLASMIC (POTENTIAL).

FT TRANSMEM 346 366 10 (POTENTIAL).

FT DOMAIN 367 384 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 385 405 11 (POTENTIAL).

FT DOMAIN 406 410 PERIPLASMIC (POTENTIAL).

FT TRANSMEM 411 431 12 (POTENTIAL).

FT DOMAIN 432 440 CYTOPLASMIC (POTENTIAL).

SO SEQUENCE 440 AA; 47208 MW; 3A23302A9DCBESBE CRC64;

Query Match 4.1%; Score 117; DB 1; Length 440;

Best Local Similarity 28.4%; Pred. No. 0.24; Indels 38; Gaps 10;

Matches 48; Conservative 29; Mismatches 54;

Qy 53 MTMVLGIGPVILGVCVPLLGASDHWGRYGRRRPFIALSLGILSLFLI-PRAGWLAG 111

Db 295 MMAVIGFG-----VMVPVAGLLAD-----AFGRKSMVIIITLILPALFAPNPLG--- 341

Qy 112 LLCPPRRP-LELALLILGVLLDFCQVCHTFPEALLSLDLFRDPDHCROAYSVAFMISL 170

Db 342 -----SGNPILVFAPLGLLSLMG-----LTFGPMGALLPELF--PTEVR--YTGFSSYNV 389

Qy 171 GGCLGYLLPAIDWDTSALAPYLGTOBECLFGLLTL-IFLTCVAATLIVA 218

Db 390 ASILG-----ASVAPYIAANLQTNVGLGAVGLYLAAMAGLTLIA 428

RESULT 11

SWVA_SALTY

ID -SWVA_SALTY STANDARD; PRT; 495 AA.

AC P37594;

DT 01-OCT-1994 (Rel. 30, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Methyl viologen resistance protein smva.

GN SWVA OR STM1574.

OS Salmonella typhimurium.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OT Enterobacteriaceae; Salmonella.

OX NCBI_TaxID=602;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SLI303;

RX MEDLINE=95011654; PubMed=7926834;

RA Hongo E., Morimyo M., Mita K., Machida I., Hama-Inaba H., Tsuji H.,

RA Ichimura S., Noda Y.;

RT "The methyl viologen-resistance-encoding gene smva of Salmonella

RT typhimurium."

RL Gene 148:173-174(1994).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=LT2 / SGSC1412 / ATCC 700720;

RX MEDLINE=21534948; PubMed=11677609;

RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,

RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,

Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E., Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M., Waterston R., Wilson R.K.;
 "Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2.";
 Nature 413:852-856 (2001).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane (Potential).
 CC AS THE DRUG RESISTANCE TRANSLOCASE FAMILY (ALSO KNOWN
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 DR EMBL; D26057; BAA05055.1; -;
 DR EMBL; AE008769; AAL20492.1; -;
 DR StyGene; SG10384; smva.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR001411; TCR_TetB.
 DR PRINTS; PR01036; TCRTEB.
 KW Transmembrane; Inner membrane; Complete proteome.
 FT TRANSMEM 5 25 POTENTIAL.
 FT TRANSMEM 44 64 POTENTIAL.
 FT TRANSMEM 73 93 POTENTIAL.
 FT TRANSMEM 96 116 POTENTIAL.
 FT TRANSMEM 135 155 POTENTIAL.
 FT TRANSMEM 158 178 POTENTIAL.
 FT TRANSMEM 192 212 POTENTIAL.
 FT TRANSMEM 220 240 POTENTIAL.
 FT TRANSMEM 260 280 POTENTIAL.
 FT TRANSMEM 299 319 POTENTIAL.
 FT TRANSMEM 327 347 POTENTIAL.
 FT TRANSMEM 357 377 POTENTIAL.
 FT TRANSMEM 391 411 POTENTIAL.
 FT TRANSMEM 469 489 POTENTIAL.
 FT TRANSMEM 444 444 A -> R (IN REF. 1).
 FT CONFLICT 147 147 G -> V (IN REF. 1).
 FT CONFLICT 182 182 V -> D (IN REF. 1).
 FT CONFLICT 198 198 A -> V (IN REF. 1).
 FT CONFLICT 227 227 F -> Y (IN REF. 1).
 FT CONFLICT 447 447 L -> Y (IN REF. 1).
 FT CONFLICT 454 457 AILD -> GKLT (IN REF. 1).
 SQ SEQUENCE 495 AA; 52134 MW; 961F77C748CDA164 CRC64;
 Query Match 4.1%; Score 117; DB 1; Length 495;
 Best Local Similarity 21.0%; Pred. No. 0.27;
 Matches 92; Conservative 60; Mismatches 151; Indels 136; Gaps 22;
 QY 17 QLLVNLTLTGL-----BVCIAAGTYVPPLLLEGV-----BEKFM 54
 DB 72 RLMLGTLFGLASLAASHTASWLTATVLLAIGAAMIVPATL-AGIRATFCEEKHN 130
 QY 55 MVLGI-----GPVLG-----LVCVPL-----LGSASDHWGRYRR- 85
 DB 131 MALGWAANVGGGAAGFLGLIGLLEHFYWGVSFLNVPVLMGLTARYVPRQGRD 190
 QY 86 RPF-----IWLSL-----GLLSLFLIPRAGMLAGL 112
 DB 191 QPLNLGHAVMLITAILLVYSKATLKGHLSLWISFTLTGALLGLFIRTLQATSRPM 250
 QY 113 LCPDPRPLEALILGVGLDFCGQVCFTEPLEALLDFDPDHCRAQSVYAFMI----- 168
 DB 251 I--DMRLFTIRILSGV-VMAWTAMITLVGFELMAQELQFV-HGLSPYEAGVFMVPMV 306
 QY 169 -----SLGCL-----GYLLPAIDNWTSAALYLTQOECLFGLTLFLTC 210
 DB 307 ASGFGPIAGVLVSRGLRLVATGGMALSALSFLGLAMTDF-STQQWQAWGLMALLGFS- 364

QY 211 VAATLLVAEEAALGTEPAE-GLSAPSLPHCCPCRRARLAFRNGLALPRHLQLCCMRP 269
 DB 365 -AASALLASTSAIMAAAPAEKAAAGAEETWAYELGAGLGAIFGLLLSRFSASIRLPA 423
 QY 270 TLRLFLVAELCSWMLMTFTFLFYTFDVGEG--LVQGVV--RAEPTGETARRH---YDEGVR 322
 DB 424 GLEAQEIARASSM-----GEAVQLANSPLPTQCAILDAAARHAFIWSHVA 470
 QY 323 MGSIG-LFLOCAISLVFSL 340
 DB 471 LSSAGSMELLALLAVGNWFSL 489
 RESULT 12
 Y051_MYCTU STANDARD; PRT; 640 AA.
 AC Q10880; O53628; (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 41, Last annotation update)
 DE 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein RV0083.
 GN RV0083 OR MT0090 OR MTCY251.01 OR MTV030.27.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=9825987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekajia P., Badcock K., Braham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Feldwell T., Gentles S., Hamlin N., Holtroyd S., Hornsby T., Jageis K., Krogh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares R., Squares R., Sultston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.";
 RT Nature 393:537-544 (1998).
 RL [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / OshKosh;
 RP Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O., Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L., Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A., Bishai W.;
 RA "Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains";
 RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 RL -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- SIMILARITY: TO POLYPEPTIDE 5 OF THE NADH-UBIQUINOL OXIDOREDUCTASE OF CHLOROPLASTS OR MITOCHONDRIA.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AL021428; CAAL6264.1; -;
 DR EMBL; AE006920; AAK44315.1; -;
 DR PIR; D70850; D70850.
 DR TIGR; MT0090; -;
 DR Tuberculist; RV0083; -;
 DR InterPro; IPR003918; NADH_oxred4.
 DR InterPro; IPR001750; Oxidored_g1.
 DR Pfam; PF00361; Oxidored_g1; 1-
 DR PRINTS; PR01437; NUOXDRDTASE4.

KW Hypothetical protein; Oxidoreductase; Transmembrane;
KW Complete proteome.
FT TRANSMEM 8 28 POTENTIAL.
FT TRANSMEM 52 72 POTENTIAL.
FT TRANSMEM 90 110 POTENTIAL.
FT TRANSMEM 136 156 POTENTIAL.
FT TRANSMEM 179 199 POTENTIAL.
FT TRANSMEM 208 228 POTENTIAL.
FT TRANSMEM 241 261 POTENTIAL.
FT TRANSMEM 277 297 POTENTIAL.
FT TRANSMEM 298 318 POTENTIAL.
FT TRANSMEM 352 372 POTENTIAL.
FT TRANSMEM 391 411 POTENTIAL.
FT TRANSMEM 446 466 POTENTIAL.
FT TRANSMEM 497 517 POTENTIAL.
FT TRANSMEM 619 639 POTENTIAL.
SQ SEQUENCE 640 AA; 65627 MW; 6254C289DBD108A8 CRC64;

Query Match 3.9%; Score 111.5; DB 1; Length 640;
Best Local Similarity 22.3%; Pred. No. 0.87;
Matches 129; Conservative 73; Mismatches 222; Indels 155; Gaps 29;

Qy 16 AQLLV---NLTFTGL--EVCLAAGITYV-----PELLLEVGVEEKFTMTWVLGIGPVLG 64
Db 93 AAMLLVPAAGSVTFLAWELMAJASLVLSEHARPOVRSAGLWYAVMTQGFIALVVG 152
Qy 65 LVCVPLGASDHWGRYGRRRPPPIWALSGLLSLFP-----LIPRAGMLAGL 112
Db 153 LVVLAAG-GSDRPA-----LGAVCDGVRAAVFMTLVGFGKAGLVPLHAWLPRA 203
Qy 113 LCPDPR-----LELALLGV-GLLDPCGGVCFPLEALLSDLPRDPDHCRCQAVSVAFM 167
Db 204 HPEAPSPVSAALMAAMVNLGIYGVRFDLQ-----LGPGRRWGLA-----L 246
Qy 168 ISLGG---CLGYLLPATDWDTSALAPYLTQEECLFGLTLTFLTCVAATLLVAEEAALG 224
Db 247 LAVGTSALYGVLOASVAADUKRLAYSTTN---MGLITL-----ALGAATLFDATGAYG 299
Qy 225 PTEPAEGISAPSLPHCCPCRCARLAFRNGLALLPR-----LHQLCCRMPTLRRLFV 276
Db 300 PASTAAAAAMLMHTAHA--AFKSLAFVAAGSVLAATGLRDLGLGLARRMPAT----- 351
Qy 277 AELCSWALMTFTLFY-----TVFFGVAAAGACGLPLGAGFVSEWLLVQSLIHAAPGHPPIVALTTPLA 399
Db 352 -----TVFFGVAAAGACGLPLGAGFVSEWLLVQSLIHAAPGHPPIVALTTPLA 399
Qy 323 MGSIGLFLQCAISLVFS-----LVMDRLVQRFCTRAVYLASVAAPFAAGATCLSHVA 376
Db 400 VGVVALATGLSVAAAMTKAFGIGFLARPRSTQAEAAAREAPASMRAGMAIAGA-CLVLAVA 458
Qy 377 --VVTASAALTGTFFSALQILPYT--LASLYHREKQVFLPKYRGDTGGASSEDLSMTSFLP 433
Db 459 PLLVAPVRRRAATLPAQAVKFTGLGAV-----VRLPAMSG-----SIAP 499
Qy 434 GPKFAGPFPNGHVAGGSGLLP-----PPPA-----LCASACDVSVRVVVGEPPEAR 481
Db 500 GVIAAAVL-----AAALAVAVLARWRRRRRPPAPARLPLWACGAA--DLTVRMQYATTSFAE 553
Qy 482 VVPGRGICLDLAILDSALFLLSQVAPSLFMSIVQLSQSV 520
Db 554 --PLQRFVGVLDPRDPTDIEVHTAESRYMAERTYRTAV 590

RESULT 13
PHDK_NOCCK
ID PHDK_NOCCK STANDARD; PRT; 473 AA.
AC 024723;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable 1-hydroxy-2-naphthoate transporter.
OS Nocardioides sp. (strain KP7).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Propionibacteriaceae; Nocardioidaceae; Nocardioides.
OX NCBI_TaxID=35761;
RN SEQUENCE FROM N.A.
RX MEDLINE=97474276; PubMed=9335300;
RA Iwabuchi T., Harayama S.;
RT "Biochemical and genetic characterization of 2-carboxybenzaldehyde
dehydrogenase, an enzyme involved in phenanthrene degradation by
Nocardioides sp. strain KP7.";
RL J. Bacteriol. 179:6488-6494 (1997).
CC -!- FUNCTION: PROBABLY INVOLVED IN THE TRANSPORT OF 1-HYDROXY-2-
NAPHTHOATE.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
(Potential).
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AB000735; BAA23264.1; --
DR InterPro; IPR007114; NPS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; sugar tr; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; FALSE_NEG.
KW Transport; Transmembrane; Inner membrane.
FT DOMAIN 1 25 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 26 46 1 (POTENTIAL).
FT DOMAIN 47 59 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 60 80 2 (POTENTIAL).
FT DOMAIN 81 91 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 92 112 3 (POTENTIAL).
FT DOMAIN 113 121 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 122 142 4 (POTENTIAL).
FT DOMAIN 143 153 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 154 174 5 (POTENTIAL).
FT DOMAIN 175 180 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 181 201 6 (POTENTIAL).
FT DOMAIN 202 263 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 264 284 7 (POTENTIAL).
FT DOMAIN 285 304 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 305 325 8 (POTENTIAL).
FT DOMAIN 326 330 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 331 351 9 (POTENTIAL).
FT DOMAIN 352 354 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 355 375 10 (POTENTIAL).
FT DOMAIN 376 398 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 399 418 11 (POTENTIAL).
FT DOMAIN 419 421 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 422 444 12 (POTENTIAL).
FT DOMAIN 445 473 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 473 AA; 49109 MW; D6D765D376260D8A CRC64;

Query Match 3.9%; Score 111; DB 1; Length 473;
Best Local Similarity 24.3%; Pred. No. 0.69;
Matches 114; Conservative 66; Mismatches 180; Indels 110; Gaps 27;

Qy 12 RHRKAQLLVNLTFT--GLEVCIAAGITYVPPLLLEVGVEEKFTMTWVLGIPVGLVCVP 69
Db 22 RRQRIIVLVNLTFT--GLEVCIAAGITYVPPLLLEVGVEEKFTMTWVLGIPVGLVCVP 69
Qy 70 LLGASDHWGRYGRRRPPPIWALSGLLSLFPILPRAG-WIAGLLCPDPPPLEALLILG 128
Db 80 VSGFVADRW-GRKG-----VTWVGFV--LFCIATAGLGTG---DIHSFAALRIISC 125
Qy 129 VGLLDPCGGVCFPLEALLSDLPRDPDHCRC-QAVSVVAFVMSLGGCLGYLLPAIDWDTS 187

Db 126 FGL-----GAVNPVAL-TIVADWM--PKARRAQMVSIAFAGVGVSGSIIGAYLAA-----A 172

Qy 188 LAPVILGTQECLEF-GLLTILFICVAAATLLVAEEA-----ALGPTEPAEG 231

Db 173 VIPILGWQVMVLIAGLAPLILPFVA--LVPEPAIISVRGIPPEARIRSRALALVAPDRD 230

Qy 232 LSAPLSLPHCCPCRARLAFRNLGALLPRHLQCCRMPTLRLRFLVAELCSWMLWTFLLF 291

Db 231 IAGVDLT-----RAGL-----TLGAGEVR-----AKALFAELICRPLLGVTLLIW 270

Qy 292 YTDVFGEG-----LYQGVF-----RAEPCGEARRHYDEGVMSGLGLFLOCAISLVFSLVMD 343

Db 271 GVFFVQSGSLVLVQIYMPMLLQAPAPGLST---VESGLIVAYMGW-----ALIGQLTIA 322

Qy 344 RLVRQFGFRVAVLASVAAPPAAGATCLSHSVAVVTGAAALTGTF-----SALQI-L 395

Db 323 FILKRF-DRFTALAAFIWSV-----VGLLIVAAFGTGFYFILLFAIGLSL 371

Qy 396 PYTLASLVHREKQVFLPKYR-----GDTGGASSEDLSMTSFLPGPKPGAPF 441

Db 372 PATAAAMQSVTTLAYEEFRATGMSAGFAGRLGLTLYGALGGTLIGAGF 421

RESULT 14

CAIA HUMAN

ID CAIA HUMAN STANDARD; PRT; 680 AA.

AC Q03692;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Collagen alpha 1(X) chain precursor.

GN COL10A1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=92109659; PubMed=1764025;

RA Thomas J.T., Creswell C.J., Rash B., Nicolai H., Jones T.,

RA Solomon E., Grant M.E., Boot-Handford R.P.;

RT "The human collagen X gene. Complete primary translated sequence and

RT chromosomal localization.";

RL Biochem. J. 280:617-623(1991).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=93012005; PubMed=1397333;

RA Reichenberger E., Beier F., Luvalle P., Olsen B.R., von der Mark K.,

RA Bertling W.M.;

RT "Genomic organization and full-length cDNA sequence of human collagen

RT X.";

RL FEBS Lett. 311:305-310(1992).

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=92267014; PubMed=1587271;

RA Apte S.S., Seldin M.F., Hayashi M., Olsen B.R.;

RT "Cloning of the human and mouse type X collagen genes and mapping of

RT the mouse type X collagen gene to chromosome 10.";

RL Eur. J. Biochem. 206:217-224(1992).

RN [6]

RP SEQUENCE FROM N.A.

RX MEDLINE=91243838; PubMed=2037056;

RA Apte S., Mattei M.-G., Olsen B.R.;

RT "Cloning of human alpha 1(X) collagen DNA and localization of the

RT COL10A1 gene to the q21-q22 region of human chromosome 6.";

RL

FEBS Lett. 282:393-396(1991).

[7]

RP SEQUENCE OF 547-655 FROM N.A.

RX MEDLINE=92077285; PubMed=1743401;

RA Reichenberger E., Aigner T., von der Mark K., Stoeb H., Bertling W.;

RT "In situ hybridization studies on the expression of type X collagen

RT in fetal human cartilage.";

RL Dev. Biol. 148:562-572(1991).

RN [8]

RP REVIEW ON VARIANTS.

RX MEDLINE=97255959; PubMed=9101290;

RA Kuivaniemi H., Tromp G., Prockop D.J.;

RT "Mutations in fibrillar collagens (types I, II, III, and XI), fibril-

RT associated collagen (type IX), and network-forming collagen (type X)

RT cause a spectrum of diseases of bone, cartilage, and blood vessels.";

RL Hum. Mutat. 9:300-315(1997).

RN [9]

RP VARIANTS SMCD ASP-598 AND PRO-614.

RX MEDLINE=94136476; PubMed=8304336;

RA Wallis G.A., Rash B., Sweetman W.A., Thomas J.T., Super M., Evans G.,

RA Grant M.E., Boot-Handford R.P.;

RT "Amino acid substitutions of conserved residues in the

RT carboxyl-terminal domain of the alpha 1(X) chain of type X collagen

RT occur in two unrelated families with metaphyseal chondrodysplasia

RT type Schmid.";

RL Am. J. Hum. Genet. 54:169-178(1994).

RN [10]

RP VARIANT SMCD ARG-591.

RX MEDLINE=94272470; PubMed=8004099;

RA McIntosh I., Abbott M.H., Warman M.L., Olsen B.R., Francomano C.A.;

RT "Additional mutations of type X collagen confirm COL10A1 as the

RT Schmid metaphyseal chondrodysplasia locus.";

RL Hum. Mol. Genet. 3:303-307(1994).

RN [11]

RP VARIANT SMCD VAL-618.

RX MEDLINE=95181449; PubMed=7876225;

RA Chan D., Cole W.G., Rogers J.G., Bateman J.F.;

RT "Type X collagen multimer assembly in vitro is prevented by a Gly618

RT to Val mutation in the alpha 1(X) NCI domain resulting in Schmid

RT metaphyseal chondrodysplasia.";

RL J. Biol. Chem. 270:4558-4562(1995).

RN [12]

RP VARIANTS SMCD ARG-545; GLU-595; HIS-597; LYS-617; ARG-644 AND GLY-648.

RX MEDLINE=95331767; PubMed=7607655;

RA Bonaventure J., Chaminade F., Maroteaux P.;

RT "Mutations in three subdomains of the carboxy-terminal region of

RT collagen type X account for most of the Schmid metaphyseal

RT dysplasias.";

RL Hum. Genet. 96:58-64(1995).

RN [13]

RP VARIANT SMCD PRO-600.

RX MEDLINE=96375754; PubMed=8782043;

RA Wallis G.A., Rash B., Sykes B., Bonaventure J., Maroteaux P.,

RA Zabel B., Wynne-Davies R., Grant M.E., Boot-Handford R.P.;

RT "Mutations within the gene encoding the alpha 1 (X) chain of type X

RT collagen (COL10A1) cause metaphyseal chondrodysplasia type Schmid but

RT not several other forms of metaphyseal chondrodysplasia.";

RL J. Med. Genet. 33:450-457(1996).

RN [14]

RP VARIANTS SMCD GLU-18 AND ARG-18.

RX MEDLINE=97220591; PubMed=9067753;

RA Ikegawa S., Nakamura K., Nagano A., Haga N., Nakamura Y.;

RT "Mutations in the N-terminal globular domain of the type X collagen

RT gene (COL10A1) in patients with Schmid metaphyseal

RT chondrodysplasia.";

RL Hum. Mutat. 9:131-135(1997).

RN [15]

RP VARIANTS SMD GLU-595.

RX MEDLINE=99057503; PubMed=9837818;

RA Ikegawa S., Nishimura G., Nagai T., Hasegawa T., Ohashi H.,

RA Nakamura Y.;

RT "Mutation of the type X collagen gene 'COL10A1' causes

RT spondylometaphyseal dysplasia.";

RL

RP SEQUENCE FROM N.A.
 RX MEDLINE=20057165; PubMed=10591208;
 RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
 RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
 RA Baguley C., Bailey J., Barlow K.P., Bates K.N., Beasley O.P.,
 RA Bird C.P., Blakey S.E., Bridgman A.M., Buck D., Burgess J.,
 RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
 RA Clagg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
 RA Dhani P.D., Dockree C., Dowlath S.J., Cox A.V., Davis J., Dawson E.,
 RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,
 RA Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,
 RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,
 RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,
 RA Laird G.K., Langford C.F., Levergha M.A., Lloyd C., Lloyd D.M.,
 RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., McCann O.J.,
 RA Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
 RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
 RA Phillips S.H., Plumb R.W., Ramsey H., Ramsey Y., Rogers L., Ross M.T.,
 RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
 RA Soudernlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,
 RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
 RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
 RA Wright C.B., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
 RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
 RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,
 RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
 RA Dorman A., Fang F., Fu Y., Hua A., Kenton S., Lai H., Lao H.I.,
 RA Lewis S., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,
 RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
 RA Zhang M., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
 RA Zhang Q., Zhang G., Chisoe S., Murray J., Miller N., Minx P.,
 RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
 RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
 RA Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,
 RA Scheet P., Walker C., Wamsley A., Wohlmann P., Pepin K., Nelson J.,
 RA Korff I., Bedell J.A., Hillier L., Mardis E., Waterston R., Wilson R.,
 RA Emanuel B.S., Shaikh T., Kurahashi H., Saitta S., Budarf M.L.,
 RA McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E., Edelman L.,
 RA Kim U.J., Shizuya H., Simon M.I., Dumanski J.P., Peyrard M., Kedra D.,
 RA Seroussi E., Franssen I., Tapia I., Bruder C.E., O'Brien K.P.,
 RA Wilkinson P., Bodenteich A., Hartman K., Hu X., Khan A.S., Lane L.,
 RA Tiahun Y., Wright H.;
 RT "The DNA sequence of human chromosome 22.";
 RL Nature 402:489-495(1999).
 CC -1- FUNCTION: RECEPTOR FOR THE HORMONE GALANIN.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF073799; AAC35944.1; -;
 CC EMBL: AF067733; AAC18860.1; -;
 CC EMBL: AF129514; AAD47348.2; -;
 CC EMBL: AF129513; AAD47348.2; JOINED.
 CC EMBL: Z97630; -; NOT ANNOTATED_CDS.
 CC EMBL: HGNC:4134; GAI3.
 CC MIM: 603692; -;
 CC GO: GO:0016021; C: integral to membrane; TAS.
 CC GO: GO:0005886; C: plasma membrane; TAS.
 CC GO: GO:0004966; F: galanin receptor activity; TAS.
 CC GO: GO:0007631; F: feeding behavior; TAS.
 CC GO: GO:0007611; P: learning and/or memory; TAS.
 CC GO: GO:0007194; P: negative regulation of adenylate cyclase ac. . . ; TAS.
 CC GO: GO:0007218; P: neuro-peptide signaling pathway; TAS.
 CC GO: GO:0007268; P: synaptic transmission; TAS.
 CC InterPro: IPR000276; GPCR_Rhodopsin.

DR Pfam: PF00001; 7tm1; 1.
 DR PRINTS: PRO0237; GPCRHOPOFSN.
 DR PROSITE: PS00237; G-PROTEIN RECEPTOR F1_1; 1.
 DR PROSITE: PS00262; G-PROTEIN RECEPTOR F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Multigene family; Lipoprotein; Palmitate.
 FT DOMAIN 1 20 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 21 41 1 (POTENTIAL).
 FT DOMAIN 42 57 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 58 78 2 (POTENTIAL).
 FT DOMAIN 79 96 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 97 118 3 (POTENTIAL).
 FT DOMAIN 119 138 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 139 159 4 (POTENTIAL).
 FT DOMAIN 160 184 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 185 205 5 (POTENTIAL).
 FT DOMAIN 206 236 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 237 257 6 (POTENTIAL).
 FT DOMAIN 258 280 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 281 368 7 (POTENTIAL).
 FT CARBOHYD 6 6 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFID 95 172 BY SIMILARITY.
 FT LIPID 308 308 PALMITATE (BY SIMILARITY).
 FT DOMAIN 218 225 POLY-ALA.
 FT DOMAIN 310 318 POLY-ARG.
 SQ SEQUENCE 368 AA; 39573 MW; E8F46B66422C3150 CRC64;
 Query Match 3.8%; Score 110; DB 1; Length 368;
 Best Local Similarity 21.9%; Pred. No. 0.63;
 Matches 94; Conservative 35; Mismatches 145; Indels 156; Gaps 17;
 Qy 87 PFIWALSGLISLLFLIPRAG---WLAGLCPDPRPLE-----LALLILGVGLDFCCQV 138
 Db 21 PVVPAL-----IFLGTGVNGLVLAIVLQPGSAWQEPGSTTDLFILNLAVALDFIL 73
 Qy 139 CFTFLEALLSDLFRDPDHCQAVSVYAFMISLGGCLGYLLPAIDWD7SALAPYLGTQEEC 198
 Db 74 CCVPFQATI-----YTLDA--W----- 88
 Qy 199 LFGLLT-----LIFLTCVAATLLVAEEAALGPTPEAGLSAPSLSPHCCPCRLAFRN 252
 Db 89 LFGALVCKAVHLILYLTMYASFT-----LAAVSVDRY-----LAVR- 125
 Qy 253 LGALLPRHLQCCRMPTTLRLFVAELCSNMALMTETLFTVDFVGEGLYQGVPRAPGCTE 312
 Db 126 -----HPLRSALRTPRNARAAGLVLLAALFSAPLYLSYGTVRYGALELCVPANE 177
 Qy 313 ARHYDEGVRMGSLGLFLQCA-ISLVFSLVMDRLVQRF-----TRAVYLASVAA 361
 Db 178 DARRRALDVATFAAGYLLPVAVVSLAVGRTIRFLWAAGVAGAGAAAEARRRATGRAGRAM 237
 Qy 362 PPVAA-----GATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLY 403
 Db 238 LAVAALYALCWGPHHALICFWYGRFAFSPATYACRLASHCLAVANSCLNPLVVALASRH 297
 Qy 404 -----HREKQVLPKVRGDTGGASSDSLTSTLPGPKGAPFNGHV 447
 Db 298 FRARFRLWPCGRRRRHARRA-LRRVRPASSG-----PPGCPGDARPSGRLL 344
 Qy 448 AGGSGLLPPP 457
 Db 345 AGG-GQPEP 353

Search completed: January 25, 2004, 17:05:24
 Job time : 31 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2004, 17:00:36 ; Search time 40 Seconds
(without alignments)
3567.577 Million cell updates/sec

Title: US-09-593-793a-113

Perfect score: 2861

Sequence: 1 MVQRLWVSRLLRRKKAQLL.....AIYFATQVDFKSLAKYSA 553

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2861	100.0	553	4 Q96JT2	Q96jt2 homo sapien
2	2798	97.8	553	6 Q95KI5	Q95ki5 macaca fasc
3	2602	90.9	553	11 Q8K0H7	Q8k0h7 mus musculus
4	2531	88.5	501	6 Q95KC5	Q95kc5 macaca fasc
5	2099	73.4	450	11 Q8K252	Q8k252 mus musculus
6	519.5	18.2	599	5 Q9SVI1	Q9svi1 drosophila
7	512	17.9	112	11 Q8RI10	Q8ri10 mus musculus
8	484.5	16.9	530	11 Q8C204	Q8c204 mus musculus
9	456	15.9	576	13 Q90274	Q90274 oryzias lat
10	446	15.6	751	11 Q8K4S3	Q8k4s3 rattus norv
11	394.5	13.8	754	5 Q8MT99	Q8mt9 dictyosteli
12	347.5	12.1	515	10 Q65803	Q65803 daucus caro
13	347.5	12.1	515	10 Q9FNR6	Q9fnr6 daucus caro
14	346.5	12.1	501	10 Q9SQK6	Q9sqk6 vitis vinif
15	346.5	12.1	515	10 Q9SLN7	Q9sln7 daucus caro
16	346	12.1	612	10 Q9SQK5	Q9sqk5 vitis vinif

17	338	11.8	512	10	Q9S7Z5	Q9s7z5 apium grave
18	337.5	11.8	523	10	Q04077	Q04077 vicia faba
19	336.5	11.8	501	10	Q9SP63	Q9sp63 vitis vinif
20	333	11.6	512	10	Q9ZTB9	Q9ztb9 apium grave
21	332.5	11.6	508	10	Q8VYX3	Q8vyx3 brassica ol
22	330	11.5	516	10	Q43653	Q43653 solanum tub
23	329.5	11.5	524	10	Q9XHL6	Q9xhl6 pisum sativ
24	328.5	11.5	506	10	Q944W2	Q944w2 oryza sativ
25	328.5	11.5	594	10	Q80605	Q80605 arabidopsis
26	327.5	11.4	506	10	Q948L0	Q948l0 oryza sativ
27	325.5	11.4	604	10	Q9FVL6	Q9fvl6 lycopersico
28	324.5	11.3	502	10	Q9SP14	Q9sp14 alonsoa mer
29	324.5	11.3	512	10	Q8RWQ6	Q8rwq6 arabidopsis
30	323.5	11.3	512	10	Q39231	Q39231 arabidopsis
31	323.5	11.3	512	10	Q80550	Q80550 arabidopsis
32	317.5	11.1	507	10	Q40583	Q40583 nicotiana t
33	317.5	11.1	530	10	Q9M535	Q9m535 euphorbia e
34	316.5	11.1	500	10	Q9FVJ6	Q9fvj6 lycopersico
35	313	10.9	521	10	Q9SXM0	Q9sxm0 zea mays m
36	311	10.9	533	10	Q41152	Q41152 ricinus com
37	309.5	10.8	607	10	Q8LPM4	Q8lpm4 citrus sine
38	307.5	10.7	488	10	Q9FV92	Q9fv92 solanum tub
39	306.5	10.7	513	10	Q8VYX4	Q8vyx4 brassica ol
40	306.5	10.7	528	10	Q8LPM6	Q8lpm6 citrus sine
41	306	10.7	523	10	Q9M422	Q9m422 hordeum vul
42	305.5	10.7	501	10	Q65929	Q65929 daucus caro
43	302.5	10.6	510	10	Q9SP15	Q9sp15 asarina bar
44	302	10.6	523	10	Q8RUL3	Q8rul3 triticum ae
45	301.5	10.5	429	10	Q40167	Q40167 lycopersico

ALIGNMENTS

RESULT 1

ID Q96JT2 PRELIMINARY; PRT; 553 AA.
 AC Q96JT2;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Prostein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Prostate;
 RX MEDLINE=21139094; PubMed=11245466;
 RA Xu J., Kalos M., Stolk J.A., Zaslloff E.J., Zhang X., Houghton R.L.,
 RA Filho A.M., Notasco M., Badaro R., Reed S.G.;
 RT "Identification and characterization of prostein, a novel prostate-
 specific protein."
 RL Cancer Res. 61:1563-1568(2001).
 DR EMBL: AY033593; AAK54386.1; -
 SQ SEQUENCE 553 AA; 59322 MW; 0AFA23FBC742A667 CRC64;

Query Match 100.0%; Score 2861; DB 4; Length 553;
 Best Local Similarity 100.0%; Pred. No. 2.6e-209;
 Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVQRLWVSRLLRRKKAQLLNLITFGLVCLAGITVPVPLLEVGVEEKFTMWLGIG 60
 Db 1 MVQRLWVSRLLRRKKAQLLNLITFGLVCLAGITVPVPLLEVGVEEKFTMWLGIG 60

Qy 61 PVLGLVCVPLLGASDHWGRYGRRRFFIWLISGLLSLFLIPRAGWLGLCPDRPL 120

Db 61 PVLGLVCVPLLGASDHWGRYGRRRFFIWLISGLLSLFLIPRAGWLGLCPDRPL 120

Qy 121 ELALLILGVGLDFCGQVCFPTPLAALSDLFDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180

Db 121 ELALLILGVGLDFCGQVCFPTPLAALSDLFDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180

```

QY 181 IDWDTSLAPYLGTQBECLFGLLTLLIFLTCAATLLVAEEAALGPTEPAEGLSAPLSPH 240
DB 181 IDWDTSLAPYLGTQBECLFGLLTLLIFLTCAATLLVAEEAALGPTEPAEGLSAPLSPH 240
QY 241 CCPCARLAFRNIGALLPRHLQCCRMPTLRRLFVAELCSWMLMTFTFLFYDDFVGEGL 300
DB 241 CCPCARLAFRNIGALLPRHLQCCRMPTLRRLFVAELCSWMLMTFTFLFYDDFVGEGL 300
QY 301 YQGVPRAEPTGEARRHYDEGVRMGSGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
DB 301 YQGVPRAEPTGEARRHYDEGVRMGSGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
QY 361 AFPAAGATCLSHSVAVVTASAAATGTFPSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
DB 361 AFPAAGATCLSHSVAVVTASAAATGTFPSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
DB 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIVQLSQSVTAYMVSAAAGLGLVAIYFATQ 540
DB 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIVQLSQSVTAYMVSAAAGLGLVAIYFATQ 540
QY 541 VVFDKSLAKYSA 553
DB 541 VVFDKSLAKYSA 553

RESULT 2
Q95K15 PRELIMINARY; PRT; 553 AA.
AC Q95K15;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE Hypothetical 59.4 kDa protein.
OC Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Temporal cortex;
RA Osada N., Hida M., Kuehda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB060851; BAB46871.1; -.
KW Hypothetical protein.
SQ SEQUENCE 553 AA; 59392 MW; 0718F3A91FB3BF1E CRC64;

Query Match 97.8%; Score 2798; DB 6; Length 553;
Best Local Similarity 98.2%; Pred. No. 1.6e-204;
Matches 542; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLRRKKAQLLLVNTLTFGLVCLAGITVVPPLLEVGVEEKFTMWLGIG 60
DB 1 MVQRLWVSRLRRKKAQLLLVNTLTFGLVCLAGITVVPPLLEVGVEEKFTMWLGIG 60
QY 61 PVLGLVCPVLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
DB 61 PVLGLVCPVLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
QY 121 ELALLIIVGLLDPCGVCFPTPEALLSDLRDPDHCRCQAYSVYAFMISLGGCIGYLLPA 180
DB 121 ELALLIIVGLLDPCGVCFPTPEALLSDLRDPDHCRCQAYSVYAFMISLGGCIGYLLPA 180
QY 181 IDWDTSLAPYLGTQBECLFGLLTLLIFLTCAATLLVAEEAALGPTEPAEGLSAPLSPH 240

```

```

DB 181 IDWDTSLAPYLGTQBECLFGLLTLLIFLTCAATLLVAEEAALGPTEPAEGLSAPLSPH 240
QY 241 CCPCARLAFRNIGALLPRHLQCCRMPTLRRLFVAELCSWMLMTFTFLFYDDFVGEGL 300
DB 241 CCPCARLAFRNIGALLPRHLQCCRMPTLRRLFVAELCSWMLMTFTFLFYDDFVGEGL 300
QY 301 YQGVPRAEPTGEARRHYDEGVRMGSGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
DB 301 YQGVPRAEPTGEARRHYDEGVRMGSGLFLQCAISLVFSLVMDRLVORFGTRAVYLASVA 360
QY 361 AFPAAGATCLSHSVAVVTASAAATGTFPSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
DB 361 AFPAAGATCLSHSVAVVTASAAATGTFPSALQILPYTLASLYHREKQVFLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
DB 421 ASSEDSLMTSFLPGPKPGAPFPNGHVAGGSGLLPPPPALCGASACDVSVRVVVGEPTEA 480
QY 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIVQLSQSVTAYMVSAAAGLGLVAIYFATQ 540
DB 481 RVVPGRGICLDLAILDSAFLLSQVAPSLFMGSIVQLSQSVTAYMVSAAAGLGLVAIYFATQ 540
QY 541 VVFDKSLAKYSA 552
DB 541 VVFDKSLAKYSA 552

RESULT 3
Q8KH07 PRELIMINARY; PRT; 553 AA.
AC Q8KH07;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE Hypothetical protein (PROSTEIN homolog).
GN 2210413P12RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Urinary bladder;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; BC031381; AAH31381.1; -.
DR EMBL; AK035428; BAC29063.1; -.
DR MGI; MGI:1922082; 2210413P12RIK.
DR InterPro; IPR005805; Rieseke.
DR PROSITE; PS00200; RIESEK_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 553 AA; 59742 MW; 6A474FEA630B4B13 CRC64;

Query Match 90.9%; Score 2602; DB 11; Length 553;
Best Local Similarity 90.9%; Pred. No. 1.3e-189;
Matches 502; Conservative 14; Mismatches 36; Indels 0; Gaps 0;

QY 1 MVQRLWVSRLRRKKAQLLLVNTLTFGLVCLAGITVVPPLLEVGVEEKFTMWLGIG 60
DB 1 MIQRLWASRLRRKKAQLLLVNTLTFGLVCLAGITVVPPLLEVGVEEKFTMWLGIG 60
QY 61 PVLGLVCPVLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
DB 61 PVLGLVCPVLGASDHWGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120

```

QY 121 ELALLILGVLLDFCGQVCFPLEALLSDFRDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
 DB 121 ELALLILGVLLDFCGQVCFPLEALLSDFRDPDHCRCQAYSVYAFMISLGGCLGYLLPA 180
 QY 181 IDWTSALAPYLQTEECLEGLLTLILTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
 DB 181 IDWTSALAPYLQTEECLEGLLTLILTCVAATLLVAEEAALGPTPEAGLSAPLSPH 240
 QY 241 CCPCRLAFRNGLGALLPRLHQLCCRMPTLRLFLVAELCSWMAALMTFTLFTVDFVGBGL 300
 DB 241 CCPCRLAFRNGLGALLPRLHQLCCRMPTLRLFLVAELCSWMAALMTFTLFTVDFVGBGL 300
 QY 301 YQGVPRAPGTEARRHYDEGVMSGLFLQCAISLVSFLVMDRLVQRFTRAVLASVA 360
 DB 301 YQGVPRAPGTEARRHYDEGVMSGLFLQCAISLVSFLVMDRLVQRFTRAVLASVA 360
 QY 361 APFVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKRGDTGG 420
 DB 361 TFPVAAAATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLPKRGDTGG 420
 QY 421 ASSDSLSMTSFLPGPKGAPPNGHVAGGSGLLPPPPALCGASACDVSVVVGEPTEA 480
 DB 421 SSGDSQTSFLPGPKGAPPNGHVAGGSGLLPPPPALCGASACDVSVVVGEPTEA 480
 QY 481 RVPFGRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAIYFATQ 540
 DB 481 RVPFGRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGLVAIYFATQ 540
 QY 541 VVFDKSLAKYS 552
 DB 541 VVFDKSLAKYS 552

RESULT 4
 Q95KCS PRELIMINARY; PRT; 501 AA.
 ID Q95KCS
 AC Q95KCS
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Hypothetical 53.4 kDa protein.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Medulla oblongata;
 RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
 RA Suzuki Y., Sugano S., Hashimoto K.,
 RT "Isolation of full-length cDNA clones from macaque brain cDNA
 RT libraries."
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB062977; BAB60745.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 501 AA; 53447 MW; 8C554BBD04EE0470 CRC64;

Query Match 88.5%; Score 2531; DB 6; Length 501;
 Best Local Similarity 97.8%; Pred. No. 2.8e-184;
 Matches 489; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 53 MTWVLGIPVLGVCVPLLGASDHWGRYGRRRPFFIWLALSLGLLSLFLIPRAGWLAGL 112
 DB 1 MTWVLGIPVLGVCVPLLGASDHWGRYGRRRPFFIWLALSLGLLSLFLIPRAGWLAGL 60
 QY 113 LCPDPRLEALLILGVLLDFCGQVCFPLEALLSDFRDPDHCRCQAYSVYAFMISLGG 172
 DB 61 LCPDPRLEALLILGVLLDFCGQVCFPLEALLSDFRDPDHCRCQAYSVYAFMISLGG 120
 QY 173 CLGYLLPAIDWDTSALAPYLQTEECLEGLLTLILTCVAATLLVAEEAALGPTPEAGL 232

DB 121 CLGYLLPAIDWDTSALAPYLQTEECLEGLLTLILTCVAATLLVAEEAALGPTPEAGL 180
 QY 233 SAPSLPHCCPCRLAFRNGLGALLPRLHQLCCRMPTLRLFLVAELCSWMAALMTFTLFTY 292
 DB 181 SAPSLPHCCPCRLAFRNGLGALLPRLHQLCCRMPTLRLFLVAELCSWMAALMTFTLFTY 240
 QY 293 TDFVGBGLYQGVPRAPGTEARRHYDEGVMSGLFLQCAISLVSFLVMDRLVQRFTR 352
 DB 241 TDFVGBGLYQGVPRAPGTEARRHYDEGVMSGLFLQCAISLVSFLVMDRLVQRFTR 300
 QY 353 AVYLASVAAPFVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLP 412
 DB 301 AVYLASVAAPFVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLYHREKQVFLP 360
 QY 413 KYRGDTGASSEDLSMTSFLPGPKGAPPNGHVAGGSGLLPPPPALCGASACDVSVRV 472
 DB 361 KYRGDTGASSEDLSMTSFLPGPKGAPPNGHVAGGSGLLPPPPALCGASACDVSVRV 420
 QY 473 VVGPTTEARVVGPRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGL 532
 DB 421 VVGPTTEARVVGPRGICLDLAILDSAFLLSQVAPSLFMGSIIVQLSQSVTAYMVSAAGLGL 480
 QY 533 VAIYFATQVVFDDKSLAKYS 552
 DB 481 VAIYFATQVVFDDKSLAKYS 500

RESULT 5
 Q8K252 PRELIMINARY; PRT; 450 AA.
 ID Q8K252
 AC Q8K252
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein (Fragment).
 GN 2210413P12RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC034084; AAH34084.1; -;
 DR MGI; MGI:1922082; 2210413P12Rik.
 DR InterPro; IPR005805; Rheske.
 DR PROSITE; PS00200; RIESKE_2; 1.
 KW Hypothetical protein.
 FT NON_TER
 SQ SEQUENCE 450 AA; 48261 MW; 72C703AEC131302C CRC64;

Query Match 73.4%; Score 2099; DB 11; Length 450;
 Best Local Similarity 90.0%; Pred. No. 1.7e-151;
 Matches 404; Conservative 12; Mismatches 33; Indels 0; Gaps 0;

QY 104 PRAGWLAGLCPDPRLEALLILGVLLDFCGQVCFPLEALLSDFRDPDHCRCQAYSV 163
 DB 1 PRAGWLAGLCPDPRLEALLILGVLLDFCGQVCFPLEALLSDFRDPDHCRCQAYSV 60
 QY 164 YAFMISLGGCLGYLLPAIDWDTSALAPYLQTEECLEGLLTLILTCVAATLLVAEEAAL 223
 DB 61 YAFMISLGGCLGYLLPAIDWDTSALAPYLQTEECLEGLLTLILTCVAATLLVAEEAAL 120
 QY 224 GPTEPAGLSAPLSHCCPCRLAFRNGLGALLPRLHQLCCRMPTLRLFLVAELCSWMA 283
 DB 121 GPTEPAGLSAPLSHCCPCRLAFRNGLGALLPRLHQLCCRMPTLRLFLVAELCSWMA 180
 QY 284 ALMTFTLFTYDFVGBGLYQGVPRAPGTEARRHYDEGVMSGLFLQCAISLVSFLVMD 343
 DB 181 ALMTFTLFTYDFVGBGLYQGVPRAPGTEARRHYDEGVMSGLFLQCAISLVSFLVMD 240
 QY 344 RLQVFCFTRAVLASVAAPFVAAGATCLSHSVAVVTASAAALTGFTFSALQILPYTLASLY 403

Db	241	REVKQFGRSVYLAASVMTFPVAAATCLSHSVVVVTAASALTGFTTSAQLILPYTLASL	300
Qy	404	HKREKQVFLPKYRGDTGGASSSDSLMTWFLPGPKPGAPFPNGHVGAGGSGLLPPPPALCGA	463
Db	301	HKREKQVFLPKYRGDAGGSGSDSQTTSLPGPKPGALFPNGHVGSGSGLLAPPPALCGA	360
Qy	464	SACDYSVRVVGEPTPEARVWPGRGICLDLALDSAFLLSQVAPSLFMGSIQVLSQSVTAY	523
Db	361	SACDYSMRVVGEPEEARVVTGRGICLDLALDSAFLLSQVAPSLFMGSIQVLSHSVTAY	420
Qy	524	MVSAAGLGLVAIFYATQVWFKSLDAKYS	552
Db	421	MVSAAGLGLVAIFYATQVVPDKNDLAKYS	449
RESULT 6			
Q9VSV1			
ID	Q9VSV1	PRELIMINARY;	PRT; 599 AA.
AC	Q9VSV1;		
DT	01-MAY-2000 (TrEMBLrel. 13, Created)		
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)		
DT	01-JUN-2000 (TrEMBLrel. 14, Last annotation update)		
DE	CG4484 protein.		
DE	CG4484.		
GN	CG4484.		
OS	Drosophila melanogaster (Fruit fly).		
OS	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;		
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophila.		
OX	NCBI_TaxID=7227;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BERKELEY;		
RX	MEDLINE=20196006; PubMed=10731132;		
RX	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,		
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,		
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,		
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,		
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,		
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,		
RA	Abri J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,		
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,		
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,		
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,		
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,		
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,		
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,		
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,		
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,		
RA	Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,		
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,		
RA	Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,		
RA	Hostin D., Houshou K.A., Howland T.J., Wei M.-H., Ibegwan C.,		
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,		
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,		
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,		
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,		
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,		
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,		
RA	Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,		
RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,		
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,		
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,		
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,		
RA	Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,		
RA	Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,		
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,		
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,		
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,		
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;		
RT	"The genome sequence of Drosophila melanogaster";		
RT	Science 287:2185-2195(2000).		
DR	EMBL; AE003552; AAF50310.1; -		

[illegible]

```
Query Match 17.9%; Score 512; DB 11; Length 112;
Best Local Similarity 91.9%; Pred. No. 1.6e-31;
Matches 102; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 442 PNGHVAGSGLLPPPPALGASCDVSVVVVGEPTARVVPGRGICLDLALDSAFLLSQVAPSLFMSIVQLSQSVTA 501
Db 1 PNGHVAGSGSGLLAPPPALGASCDVSVVVVGEPTARVVPGRGICLDLALDSAFLL 501

Qy 502 SQVAPSLFMSIVQLSQSVTAYMVSAAGLGLVAIYFATQVVFDSKLAKYS 552
Db 61 SQVAPSLFMSIVQLSHSVTAYMVSAAGLGLVAIYFATQVVFDSKLAKYS 111

RESULT 8
ID Q8C204 PRELIMINARY; PRT; 530 AA.
AC Q8C204;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Membrane associated transporter protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK089932; BAC41003.1; -.
SQ SEQUENCE 530 AA; 57935 MW; 426898973B7D6420 CRC64;

Query Match 16.9%; Score 484.5; DB 11; Length 530;
Best Local Similarity 26.1%; Pred. No. 1.1e-28;
Matches 146; Conservative 84; Mismatches 229; Indels 101; Gaps 12;

Qy 17 QLLVNLTLFGVLCAGITVYVPLLEVGVEKFMVTLGIGVGLVCVPLGASD 76
Db 34 RLVMHSMAMGFEFCYAVEAYVTPGLLSVGLPKSLYMWLLSPILGFLQPVVGSAD 93

Qy 77 HWRGRRRRPFIWALSGLILSLFIPRAGWAGLCPDRP---LELALLIGVGLD 133
Db 94 HCRARWRRRRPYILTATMLLGMALYNGDAVSVLVANPRQKLIWAISITWVGWVLD 153

Qy 134 FCGVCFPTPLBLLSDFRDPHCRQAYSVVAFMISLGGCLGYLLPAIDWDTSALAPYL 193
Db 154 FSADFIDGPIKAYLFDVCSHQDK-EKGLHYHALFTGFGALGYLIGALDWHLDLGRLLG 212

Qy 194 TOECLFGLLTLFILTCAVALLVAEEAL--GPTPE-----AEGLSAPLSPHCCPC 244
Db 213 TEFQWMPFSSALVLLICFITHLCSIPAPLDAATDPPSQDQPGSSLSASGMHEY---- 268

Qy 245 RARLAFRNLA-----LLPRHLQCCRMPTRLRLFVAELCSW 282
Db 269 GSIEKVKNGGADTEQVQEWKNQKPSGQSQRTSMKSLLRALVNMPSHYRCLCVSHLIG 328

Qy 283 MALMTFTLYTDFVGEGLYQGVPAEPTGTEARRHYDEGVMSGLGLFQCAISLVFSLVM 342
Db 329 TAFLSNMLFFTFDMGQIVYHGDYVGAHNSFTFLYERGVGCGWGLCINVSFSSVSYFQ 388

Qy 343 DLVQRFCTRAVYLASVAFFVAAGATCLSHSVAVVTTASALTGFTFSALQILPYTLASL 402
Db 389 KAMVSYIGLGLYFMGYLLFGLGTGFIGLFNNVSTLVLCMFGVMSSTLYTVFNLIAE 448

Qy 403 YHREKQVFLPKYRGDTGASSEDLSMTSFLPGPKGAPFPNGHVAGSGSLPPPPALCG 462
Db 449 YHREEE---KEGQEA-----PGGPDNQGR---GGV-----DCA 477
```

```
Qy 463 ASACDVSVVVGEPTARVVPGRGICLDLALDSAFLLSQVAPSLFMSIVQLSQSVTA 522
Db 478 ALTCMVQL-----AQILVGG-----IGFLVNMAGSVV 506

Qy 523 YMVSAAGLGLVAIYFATQV 542
Db 507 VVITASVSLGCCFVALFV 526

RESULT 9
ID Q90274 PRELIMINARY; PRT; 576 AA.
AC Q90274;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE Membrane-associated transporter protein B.
OS Oryzias latipes (Medaka fish) (Japanese ricefish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
OX NCBI_TaxID=8090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=HNI;
RX MEDLINE=21372467; PubMed=11479596;
RA Fukumachi S., Shimada A., Shima A.;
RT "Mutations in the gene encoding B, a novel transporter protein, reduce
RT melanin content in medaka."
RL Nat. Genet. 28:381-385 (2001).
DR EMBL; AF32510; AAK7024.1; -.
SQ SEQUENCE 576 AA; 63411 MW; 13A5C982ABEB894 CRC64;

Query Match 15.9%; Score 456; DB 13; Length 576;
Best Local Similarity 24.0%; Pred. No. 1.7e-28;
Matches 144; Conservative 95; Mismatches 198; Indels 164; Gaps 17;

Qy 12 RHRKAQILLNLTLFGVLCAGITVYVPLLEVGVEKFMVTLGIGVGLVCVPLL 71
Db 56 RRSRGRILLHSMWFGREFCYAVEAAVTPVLLSVGLPRSLYSLWLISILGFLQPII 115

Qy 72 GSASDHWGRYGRRRPFIWALSGLIL-----LSLFL-----IPRAGWAGL 114
Db 116 GSASDYCRSSWGRRRPYI--LVLGILMLVGLSMFLNGDAVSVLSVDSRSSTW---- 167

Qy 115 PDRPLBLALLIGVLLDFCGQVCFTPLEALLSDL--FRDPHCRQAYSVVAFMISLGG 172
Db 168 -----AIVVMFVGVLFDFAADFDIGPIKAYLFDVGSYQDK-----RGLHYHALFTGLG 218

Qy 173 CLGYLLPAIDWDTSALAPYLGTQEECLF-----GLTLIFLTCV----- 211
Db 219 ACGYLVGAMDWGSHVGLRGLGSEYQVIFYFSALTWGVFLIVHLFSIPEKPLKVPSSSA 278

Qy 212 -AATLLVAEEA-----ALG--PTEPAEGLSAPLSPHCCPC----- 244
Db 279 SSARLLGPHSNGVGLGKEFVSPVITSSPEIPRPSYSGALGERPRFSALGEANSVTSS 338

Qy 245 -----RARLAFRNIGALLPRHLQCCRMPTRLRLFVAELCSWMLMTFTLYTDFV 296
Db 339 AKQPIKEDQKMTFRS-----LMKAIFNNPNHYRFLICISHLIGWAAFLCNMLFFTFDM 391

Qy 297 GEGYQGVPAEPTGTEARRHYDEGVMSGLGLFQCAISLVFSLVMDRLVQRFCTRAVYL 356
Db 392 GQIVYRGNPYAEHNSTAYITTYERGVGCGWGLCINVSALYSYVQVFLPYIGLGLYF 451

Qy 357 ASVAFFVAAGATCLSHSVAVVTTASALTGFTFSALQILPYTLASLYHREKQVFLPKYRG 416
Db 452 MGTVFVFGMGTSLGLFPEVIATLILCSVFGVMSSTLYTIFNLIAEQREBEQV-KLEG 510

Qy 417 DTGASSEDLSMTSFLPGPKGAPFPNGHVAGSGSLPPPPALCGASACDVSVVVVGE 476
Db 511 -----GNESPRGTGM-----DCAALTQWQL----- 531
```

```
QY 477 PTEARVVGRCICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAYI 536
Db 532 ---AQIIVGAG-----LGLVNAAGSVIVVVLSSASSISLIGCI 566
QY 537 F 537
Db 567 F 567

RESULT 10
Q8K4S3 PRELIMINARY; PRT; 751 AA.
AC Q8K4S3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE proton-associated sugar transporter A.
GN PAST-A.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus; Rattus.
OX NCB1_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Shimokawa N., Okada J., Haglund K., Dikic I., Koibuchi N., Miura M.;
RT "Past-A, a new class proton-associated sugar transport facilitator in
RL the central nervous system." to the EMBL/GenBank/DBJ databases.
DR EMBL; AB075229; BAB97313.1; -
KW Sugar transport.
SQ SEQUENCE 751 AA; 81751 MW; 5C33E1C2E4A1B2EC CRC64;

Query Match 15.6%; Score 446; DB 11; Length 751;
Best Local Similarity 22.3%; Pred. No. 1.3e-25;
Matches 160; Conservative 79; Mismatches 206; Indels 272; Gaps 15;

17 QLLVNLTLTFLGVLCAAGITYVPPLLEVGVEKFMVTLGIGPVLGLVCPVLGASD 76
86 ELLFNGCILGIEFSYAMETAYTVPLVQLQGLPQLSLVWFISPIGLFLQLPGLGWS 145
77 HWRGRRPFIWALSIGILLSLFLIPRAGWLAGLCPDP---RPLELALLIGVGLLD 133
146 RCTSRGRRRPFILVLAIGALLGLSLIN-GRDTCMALADTATNHNKNGILLTVCGVLM 204
134 PCGVCTPLEALLSLDRDPDHCROQSVYAFMISLGGCLGYLLPAIDWDTSALAYLG 193
205 PSADSADNPSHAYMWDVCGPVDQDR-GLNTHALMAGLGGGFGYVGVGHWDKTSFG 263
194 TQEBCLFGLTLFLTCVAAATLLVA-----BEAA 222
264 GOLAVIY-IFTAILSTVTFTLSIPERPLRPLGKRTAMKSPSLPSPVLLBEGA 322
223 LGPTPE---AEGLSAPLSPHCCP----- 243
323 -GDTLPSTATSLVAFSSPSIPSPSLTPKYGSFISRDSSLTGINEPASSFGTSNIDSV 381
244 -----CRALAFNLGALL----- 257
382 IDCFTAGHDNYLALPSSVPRQASVSPFAPADPGYFCQERGLERREGPLTLGLDGVLR 441
258 -----PR----- 259
442 SLDTSKPRASGILKRPQTALPDVAGNGPETSRRNVTFSSQVANILLNGKYESBLTG 501
260 -----LHQLCC---RMPRTLRLFLVAELCSWMLMTFTLYTDFVGEGLYQGVPR 306
502 SSEQSEQPLSLRLCSTIYNMPPVNLVNHFLGMLSFEGMLLFYDFMGVGVFGDPK 561
307 AEPGTEARRHYDEGVRMGSIGLFLQCAISLVFSLVMDRLVQRFQFTRAVYLASVAFVAA 366
562 APHASEAYQKYNVGTVGTCWGMCIYAFSAFYSAILEKLEBUCSVRTLTYFTAYLLFLGT 621
```

```
QY 367 GATCLSHSVAVVTASAAALTGFTSALQILRYTLASLYHREKQVLPKYRGDTGGSSEDS 426
Db 622 GLATLSRNLVYVLSLCTHYGILFTLCTLFYSLCDYQSK-----KFAG-----SSADG 671
QY 427 LMTSFLPGPKGAPFFNGHVAGSGGLLPPPPALCGASACDVSVRVVVGTEARVYVGR 486
Db 672 TR-----R 674
487 GICLDLAILDSAFLLSQVAPSLFMGSIQVLSQSVTAYMVSAAGLGLVAYIYFATQVWF 543
Db 675 GMGVDISLLSCQFLAQILVSLVGP-----LTSAVGSANGVMVFASLVSF 720

RESULT 11
Q8MMT9 PRELIMINARY; PRT; 754 AA.
AC Q8MMT9;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE SUC1-sucrose proton symporter.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCB1_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and Analysis of Chromosome 2 of Dictyostelium."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL; AC117076; AC13692.1; -
DR InterPro; IPR005828; Subtransporter.
DR Pfam; PF00083; sugar_tr; 1.
KW Transmembrane.
SQ SEQUENCE 754 AA; 82774 MW; 614C473C832ACCE4 CRC64;

Query Match 13.8%; Score 394.5; DB 5; Length 754;
Best Local Similarity 23.5%; Pred. No. 1.1e-21;
Matches 128; Conservative 90; Mismatches 200; Indels 127; Gaps 13;

QY 13 HRKA----QLLVNLTLTFLGVLCAAGITYVPPLLEVGVEKFMVTLGIGPVLGLV 68
197 HKKTLPLENLICITICFLGVQFGWALQIAFSTFLLELGVQKWSYINLAGPISGLIV 256
69 PLLGSADHWGRGRRRPFILWALSIGILLSLFLIPRAGWLAGL---CPDPRLELALLI 126
257 PLGVITDRSECRGRRKPPILIGSVFISGLVLI SNAETFGSYFGDSEQKKSIAISAI 316
127 LGVGLLDFCQVCTPLEALLSLDRDPDHCROQSVYAFMISLGGCLGYLLPAIDWDTS 186
317 VGFWLDLSNNAVQACRALLVDI-RAPSOQSLGSLFSLMLGTNLLGYMMSI--DLV 373
187 ALAPYLGTQEBCLFGLTLFLTCVAAATL-LVAEEAALGPTPEAGLSAPLSLPHCCPCR 245
374 RMVPEMKTDFRALFTLSIMVLLFCVMTLGFVTEEQYIRVNE-QSVNEF----- 422
246 ARLAFNLGALLPRLHQLCCRMPTRLRLFLVAELCSWMLMTFTLYTDFVGEGLYQGV 305
423 -----LKTMPKGIKVMPTYLQRLCAVQFFSWGFWFVLTITVWGVNVFGDP 471
306 RAEPGTEARRHYDEGVRMGSIGLFLQCAISLVFSLVMDRLVQRFQFTRAVYLASVAFV 365
472 NAFPEYSDSRILFQDGVGRWGSLSLTISGITIAVSLTLPFLVKFIDMKYITIGNML----- 526
366 AGATCLSHSVAVVTAS-----AALTGTFSSALQILPYTLASLYHREKQVLPKYRGDT 418
527 --LQCIFFALFYFVESKIGSLLLIAGTGIPIWAILPFSIVGM----- 568
419 GGASSEDLSMTSFLPDKPCAPPNGHVAGSGGLLPPPPALCGASACDVSVRVVVGTEPT 478
```



```
Db 569 -GVEDNES-----SG----- 577
Qy 479 EARVVGRCIGLIDLALDSAFLLSQVAPSLFMGSIQVLSQSVTAYMWSAAGLGLVAIYPA 538
Db 578 -----LNIQTILFIVVPPQVQVSLGILGILDLISKGNVYSILT---GSVASFFA 623
Qy 539 TQVWF 543
Db 624 TLFCF 628

RESULT 12
ID O65803 PRELIMINARY; PRT; 515 AA.
AC O65803;
DT 01-AUG-1998 (TReMBLrel. 07, Created)
DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Sucrose/H+ symporter.
GN SUT2.
OS Daucus carota (Carrot).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; campanulids; Apiales; Apiaceae; Daucus.
OX NCBI_TaxID=4039;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nantaise;
RA MEDLINE=99063785; PubMed=9847123;
RA Shaky R., Sturm A.;
RT "Characterization of source- and sink-specific sucrose/H+ symporters
from carrot.";
RL Plant Physiol. 118:1473-1480(1998).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL; Y16768; CAA76369.1; -.
DR InterPro; IPR005828; Sub transporter.
DR InterPro; IPR005989; Suc/H symport.
DR Pfam; PF00083; sugar tr; 1.
DR TIGRFAMs; TIGR01301; GPH_sucrose; 1.
KW Transmembrane.
SQ SEQUENCE 515 AA; 54424 MW; 897B90657C9E243C CRC64;

Query Match 12.1%; Score 347.5; DB 10; Length 515;
Best Local Similarity 25.9%; Pred. No. 2.7e-18;
Matches 131; Conservative 80; Mismatches 215; Indels 79; Gaps 17;
Qy 17 QLLLVNLLTFGLVCLAAAGITVPPLLELVGVVEKFTMTVLGIGVPLGLVCVPLLGASD 76
Db 34 KLVLVAIAAGVQFGWALQSLTTPYVQLLGIPIHKWAAVWLCGPISGMLVQPIVGYSD 93
Qy 77 HWRGRRRRPFIWALSIGLLSLFLIPRAGWLAGL-----LCPDPRPLEALLILGV 129
Db 94 HCQSFGRRRPFPIASGAGCAISVILI---GFAADISYKAGDDMSKTLKPRAVTVFVIGF 150
Qy 130 GLDFCGOVCFPLEALLSDLFR-DPDHCRQAYSVYAFMISLGGCLGY-----LL 178
Db 151 WILDVANMLQGPCRALLADLCSDGTRRMRSANAFYFFMAVGNILGYAAGSYNNLYKLF 210
Qy 179 PAIDWDTSALAPYLGTQBECLFGLTLTFLTCVAATLLVAEEALGPTEPAEGLSAPLS 238
Db 211 PFS---KTHACDLYCANLSCFTIISALLIITVVALSVVRENS--GPPDDADAAEEP--- 263
Qy 239 PHCCPCRLAFRNIGALLPRHQICCRMPRTLRLFVAELCSWMAWMTFTFLFYDFVGE 298
Db 264 ----PSSGKIPV--FGELLGALKDL-----PRPMLLLIIVTCLNWIAPFPFILFDTDMGR 313
Qy 299 GLYGVPAEPCTEARRHYDEGVMSGLFLQCAISLVFSLVMDRLVQRFGTAVYLAS 358
Db 314 EYGGT---AGQG---KLYDQGVRAAGLGLLNSVVLGLTIAVEYLVRGVGGVKI-LWG 366
Qy 359 VAAPFVAAGATCL-----SHSV-----AVVTASA-----ALTGFTTSALQIL 395
```

```
Db 367 FVNFILAIGLVMTVVVSKVAHQREHSANGQLPPSAGVKAGALSLSILGIPLSITYSI 426
Qy 396 PYTLASLYHREKQVFLPKYRGDTGCGASSEDLSMTSFLPGPKPG-----APFNGHVGA 448
Db 427 PFALASYSSGSGAGQGLSLGLVNLAIWVPMIVSVLAGPDSLFGGNGNLPFVVGALSA 486
Qy 449 GSGGLP-----PPALCGASACDVS 469
Db 487 AISGVLAIVLLPKPSKDAASKLSLS 511

RESULT 13
Q9FNR6
ID Q9FNR6 PRELIMINARY; PRT; 515 AA.
AC Q9FNR6;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Sucrose/proton symporter.
GN SUT2.
OS Daucus carota (Carrot).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; campanulids; Apiales; Apiaceae; Daucus.
OX NCBI_TaxID=4039;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nantaise;
RA Sturm A.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nantaise;
RA Shaky R.;
RL Thesis (2000), Department of Department of Botany,
University of Basel, Basel, Switzerland.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL; AJ303199; CAC19689.1; -.
DR InterPro; IPR005828; Sub transporter.
DR InterPro; IPR005989; Suc/H symport.
DR Pfam; PF00083; sugar tr; 1.
DR TIGRFAMs; TIGR01301; GPH_sucrose; 1.
KW Transmembrane.
SQ SEQUENCE 515 AA; 54425 MW; 9580978C726523DC CRC64;
```

```
Query Match 12.1%; Score 347.5; DB 10; Length 515;
Best Local Similarity 25.9%; Pred. No. 2.7e-18;
Matches 131; Conservative 80; Mismatches 215; Indels 79; Gaps 17;
Qy 17 QLLLVNLLTFGLVCLAAAGITVPPLLELVGVVEKFTMTVLGIGVPLGLVCVPLLGASD 76
Db 34 KLVLVAIAAGVQFGWALQSLTTPYVQLLGIPIHKWAAVWLCGPISGMLVQPIVGYSD 93
Qy 77 HWRGRRRRPFIWALSIGLLSLFLIPRAGWLAGL-----LCPDPRPLEALLILGV 129
Db 94 HCQSFGRRRPFPIASGAGCAISVILI---GFAADISYKAGDDMSKTLKPRAVTVFVIGF 150
Qy 130 GLDFCGOVCFPLEALLSDLFR-DPDHCRQAYSVYAFMISLGGCLGY-----LL 178
Db 151 WILDVANMLQGPCRALLADLCSDGTRRMRSANAFYFFMAVGNILGYAAGSYNNLYKLF 210
Qy 179 PAIDWDTSALAPYLGTQBECLFGLTLTFLTCVAATLLVAEEALGPTEPAEGLSAPLS 238
Db 211 PFS---KTHACDLYCANLSCFTIISALLIITVVALSVVRENS--GPPDDADAAEEP--- 263
Qy 239 PHCCPCRLAFRNIGALLPRHQICCRMPRTLRLFVAELCSWMAWMTFTFLFYDFVGE 298
Db 264 ----PSSGKIPV--FGELLGALKDL-----PRPMLLLIIVTCLNWIAPFPFILFDTDMGR 313
Qy 299 GLYGVPAEPCTEARRHYDEGVMSGLFLQCAISLVFSLVMDRLVQRFGTAVYLAS 358
Db 314 EYGGT---AGKG---KLYDQGVRAAGLGLLNSVVLGLTIAVEYLVRGVGGVKI-LWG 366
```

QY 359 VAAPVPAAGATCL-----SHSV-----AVVTASA-----ALTGFTSALQIL 395
 Db 367 FVNILALGLVMTVVVSKVAHQHSHSANGQLLPSAGVKAGALSLSILGIPUSIYSI 426
 QY 396 PYTLASLYHREKQVFLPKYRGDTGASSEDLSMTSFLPKPKPG-----APFPNGHVGA 448
 Db 427 PFALASIYSSGAGGGLSLGVNLAIIVVQMIYVLAGPDSLFGGNNLPFAVVGALSA 486
 QY 449 GGSGLP-----PPPALCGASACDVS 469
 Db 487 AISGVLAIVLLPKPSKDAASKLSLS 511

RESULT 14
 Q9SQK6 PRELIMINARY; PRT; 501 AA.
 AC Q9SQK6;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Putative sucrose transporter.
 GN VWSUC11.
 OS Vitis vinifera (Grape).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Vitaceae;
 OC Vitis.
 OX NCBI_TaxID=29760;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Shiraz; TISSUE=Fruit;
 RA Davies C., Wolf T., Robinson S.P.;
 RT "Three putative sucrose transporters are differentially expressed in
 grapevine tissues."
 RL Plant Sci. 147:93-100 (1999).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 DR EMBL; AF021808; AAF08329.1; -;
 DR InterPro; IPR005828; Sub.transporter.
 DR InterPro; IPR005989; Suc/H.symport.
 DR Pfam; PF00083; sugar tr; 1.
 DR TIGRFAMs; TIGR01301; GPH_sucrose; 1.
 KW Transmembrane.
 SQ SEQUENCE 501 AA; 53938 MW; 4D0D4DE2BF2F4BA8 CRC64;

Query Match 12.1%; Score 346.5; DB 10; Length 501;
 Best Local Similarity 27.0%; Pred. No. 3.1e-18;
 Matches 153; Conservative 83; Mismatches 197; Indels 133; Gaps 22;

QY 17 QLLVNLTLFGLVCLAAAGITYVPPILLLEVGVEEKFWMTVLGIGPVGLVCVPLLSASD 76
 Db 29 RLLRVASVACGIQFGWALQSLTTPYQELGIPHAWSIIWLCPGLSGLLVQPLVGHLS 88
 QY 77 HWRGRRRRPFIWALSIGLLSLFLIPRAGWLAGLL--CPDRPLELALLILGVGLDF 134
 Db 89 RCSRFRRRPFIWAGATSIWAVLIIGFSDIGGLGDGDRPRAVATFVGVFWLLDV 148
 QY 135 CGQVCFTPLEALLSDFRDPH--CROAYSVYAFMISLGGCLGYLLPA-----IDWDT 186
 Db 149 ANNVYQGPCRALLADL-TEKHRRTRVANAYFSLFIAVGNVLGATGSYSGWFRIFWFTS 207
 QY 187 ALAPYLTQBECL-----FGLTLFLTCVAATLLVAEEAALGPTPEAGLSAPSLPHCC 242
 Db 208 TSS-----CNADCANLKSAPLLDIIF---IAITTVISITAA-----QELPLSSSRSTHIS 255
 QY 243 -----PCRARLAFRNIGALLPRHLQCCRMPTLRR-----LFVAELCSNMALMTFT 289
 Db 256 EEMAETHAQEAF-----LWELP-----GTLRYLSGSIWILFTVAL-TWIGLLPFL 301
 QY 290 LFYTFDVGEGLYQGVPAEPGTEARRHYDEGVRMGSLGLFLQCAISLVFLVMDRLVQRF 349
 Db 302 LFDTDMMGREIYGGK-----NEGQNTYGVRMGALGLMNSVVLGITSVLMKLCRKW 355
 QY 350 GTRAVYLASVAAFPVAGATCLSHSVAVVTASAALTGFTSALQILPYTLASLYHREKQV 409

Db 356 GAGFVW-----CLSNILMSLCFLMLILS-----AVVKHMD--- 386
 QY 410 FLPKYRGDTGASSEDLSMTSFLPKPKGAPFPNGHVAGGSGLLPPPPALCGASACDVS 469
 Db 387 -----FLGHDLPL-----PSGVVIA-----ALIVFSILGIP 411
 QY 470 VRVVVGEV---TEARVVP-CRGICLDLAIILDSAFLLSOVAPSLPMGSIVOL-----SOSV 520
 Db 412 LAITYSVFYALISTRIESLGLGQGLSGMVLNLAIVIPQIVIVLSGSGPMDQLFGGNSPSL 471
 QY 521 TAYMVSAAAGLVAIYFATQVFPDKS 546
 Db 472 AVAAVAFASGLVAILAIPRSSADKS 497

RESULT 15
 Q9SLN7 PRELIMINARY; PRT; 515 AA.
 AC Q9SLN7;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Sucrose transporter protein.
 GN CSUT.
 OS Daucus carota (Carrot).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; Campanulids; Apiales; Apiaceae; Daucus.
 OX NCBI_TaxID=4039;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Root;
 RA Wu N., Diao F., Zhang L., Huang M.;
 RT "Sucrose transporter protein."
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 DR EMBL; AB036758; BAA89458.1; -;
 DR InterPro; IPR005828; Sub.transporter.
 DR InterPro; IPR005989; Suc/H.symport.
 DR Pfam; PF00083; sugar tr; 1.
 DR TIGRFAMs; TIGR01301; GPH_sucrose; 1.
 KW Transmembrane.
 SQ SEQUENCE 515 AA; 54523 MW; 2B548D1D9DFF51AC CRC64;

Query Match 12.1%; Score 346.5; DB 10; Length 515;
 Best Local Similarity 27.2%; Pred. No. 3.2e-18;
 Matches 137; Conservative 76; Mismatches 215; Indels 75; Gaps 19;

QY 17 QLLVNLTLFGLVCLAAAGITYVPPILLLEVGVEEKFWMTVLGIGPVGLVCVPLLSASD 76
 Db 34 KLVLVAAIAGVQFGWALQSLTTPYVQLGIPHKWAAIYWLCPISGMLVQPIVGYSD 93
 QY 77 HWRGRRRRPFIWALSIGLLSLFLIPRA--GWLAG-LLCPDRPLELALLILGVGLL 132
 Db 94 HCQSSFRRRPFIWAGATSIWAVLIIGFADIGYAGDDMSKTLKPRATVTVFVIGFWIL 153
 QY 133 DFCQVCFTPLEALLSDFR-DPDHCRQAYSVYAFMISLGGCLGY-----LLPAI 181
 Db 154 DVANNMLQGPCRALLADLCSGDTMRMSANAFYSFFMAVGNILGAAGSYNNLYKLFPPS 213
 QY 182 DWTDSALAPYLTQBECL-LFGLTLFLTCVAATLLVAEEAALGPTPEAGLSAPLSLPH 240
 Db 214 --KTHACDLYCANLKSCTIISIALIITVVA--LSVREKQWSP-DDADADEP----- 263
 QY 241 CCPCARLAFRNIGALLPRHLQCCRMPTLRLRFLVAELCSNMALMTFTLFYTFDVGEG 300
 Db 264 --PSSGKIPV--FEGELGALKDL-----PRMLLLIIVTCLNIAWPFILFDIDWGREI 315
 QY 301 YQGVPAEPGTEARRHYDEGVRMGSLGLFLQCAISLVFLVMDRLVQRFOTRAVYLASVA 360
 Db 316 YGCT--AGKG-----KLYDQGVRAAGSLGLLNSVVLGTSIAVEYLVRGVGVKI-LMGVV 368
 QY 361 AFPVAGATCL-----SHSV-----AVVTASA-----ALTGFTSALQILPY 397

Db	369	NFI	LA	GL	VM	TV	VS	KV	AQ	HR	HS	ANG	QL	PP	SAG	VK	AG	AL	SF	SL	GI	PL	SI	TY	SI	PF	428	
Qy	398	TL	AS	LY	HR	EQ	VFL	PK	YR	GD	TG	GS	SE	DS	LM	TS	FL	PG	PG	-----	AP	PN	GH	VG	AG	450		
Db	429	AL	AS	YSS	SG	AG	QGL	SL	GL	VL	LA	IV	VP	QI	VS	VL	AG	PD	SL	FG	GN	LP	AF	VG	AI	SA	AI	488
Qy	451	SGL	LP	----	PPP	AL	CG	AS	AC	DVS	469																	
Db	489	SG	VL	AI	VLL	PK	PS	KA	AS	KL	SL	SL	511															

Search completed: January 25, 2004, 17:04:44
Job time : 44 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)